

FINAL REPORT

**THE
DEVELOPMENT OF PRECISION STATEMENTS
FOR SEVERAL ASTM TEST METHODS**

**MAY 1962
NO. 13**

**Joint
Highway
Research
Project**

by
S. J. HANNA

**PURDUE UNIVERSITY
LAFAYETTE INDIANA**

THE HISTORY OF THE
CITY OF BOSTON

By
JOHN H. COOK,
Author of "The History of the City of Boston,"
"The History of the City of New York,"
"The History of the City of Philadelphia,"

and
"The History of the City of London."
In two volumes.
The first volume contains the history of the city from its
first settlement in 1630 to the year 1780. The second
volume contains the history of the city from 1780 to the
present time. The work is a complete and accurate
history of the city, and is highly recommended to all
who are interested in the history of the city.

Published by J. H. Cook, 100 N. 2nd St., Boston.

1880

1881

1882

1883

1884

1885

1886

1887

1888

1889

Final Report

THE DEVELOPMENT OF PRECISION STATEMENTS FOR SEVERAL ASTM TEST METHODS

TO: K. B. Woods, Director
Joint Highway Research Project

May 22, 1962

FROM: H. L. Michael, Associate Director
Joint Highway Research Project

File: 5-15-1
Project: C-36-65A

Attached is a final report entitled "The Development of Precision Statements for Several ASTM Test Methods". The report has been authored by Mr. Steven J. Hanna, Graduate Assistant on our staff, and was also used by Mr. Hanna as his thesis for part of the requirements of the M.S.C.E. degree. The research was conducted under the direction of Professor J. F. McLaughlin of our staff.

The objective of this research was to develop experimental designs with appropriate analyses from which precision statements (how well individual test results of a series ought to agree with each other or how well test results from different laboratories on "identical" samples ought to agree) could be developed. Several ASTM test methods were chosen and appropriate tests made. The results were statistically analyzed and precision statements for the tests formulated.

The report is presented to the Board for the record.

Respectfully submitted,

Harold L. Michael

Harold L. Michael, Secretary

HLM:kmc

Attachment

Copy:

F. L. Ashbaucher
J. R. Cooper
W. L. Dolch
W. H. Goetz
F. F. Havey
F. S. Hill
G. A. Leonards

J. F. McLaughlin
R. D. Miles
R. E. Mills
M. B. Scott
J. V. Smythe
J. L. Waling
E. J. Yoder

100
100
100
100
100

100

100

100

100
100
100
100
100
100
100
100
100
100

100
100
100
100
100
100
100
100
100
100

Final Report

THE DEVELOPMENT OF PRECISION STATEMENTS
FOR SEVERAL ASTM TEST METHODS

by

Steven J. Hanna
Research Assistant

Joint Highway Research Project
File No: 5-15-1
Project No: C-36-65A

Purdue University
Lafayette, Indiana

May 22, 1962

ACKNOWLEDGMENTS

This investigation was sponsored by the Joint Highway Research Project at Purdue University, Professor K. B. Woods, Director. The writer is grateful to this organization for providing the necessary equipment and financial support for the accomplishment of this investigation.

The writer wishes to express his sincere appreciation to his major professor and advisor, J. F. McLaughlin, for his encouragement and advice during the preparation of this thesis.

The writer wishes to express his appreciation to Professor I. W. Burr and his staff whose advice on statistical procedures was most valuable.

The writer wishes to thank Committee C-9 of the ASTM for their interest in this project.

TABLE OF CONTENTS

	Page
LIST OF TABLES	vi
LIST OF ILLUSTRATIONS	vii
ABSTRACT	viii
INTRODUCTION	1
REVIEW OF LITERATURE	6
PURPOSE AND SCOPE	14
STATISTICAL PROCEDURES	15
PROCEDURES	25
Test Method C 117-49	25
Test Method C 127-59	27
Test Method C 128-57	28
DISCUSSION OF RESULTS	33
Results of Test Method C 117-49	33
Confidence Limits for Operator A	40
Control Limits for Operator A	41
Confidence Limits for Operator B	41
Control Limits for Operator B	43
General Observations	43
Results of Test Method C 127-59	45
Absorption	48
Specific Gravity	53
General Observations	54
Results of Test Method C 128-57	54
Absorption	54
Specific Gravity	63
General Observations	64
Results Applicable to the Development of Precision	68
Statements	68
SUMMARY OF RESULTS AND CONCLUSIONS	73

Digitized by the Internet Archive
in 2011 with funding from
LYRASIS members and Sloan Foundation; Indiana Department of Transportation

TABLE OF CONTENTS (continued)

	Page
Test Method C 117-49	74
Test Method C 127-59	76
Test Method C 128-57	76
Precision Statements for Test Methods Studied	77
Test Method C 117-49	78
Test Method C 127-59	78
Test Method C 128-57	78
PROPOSED RESEARCH	79
BIBLIOGRAPHY	80
APPENDIX	82
Bartlett Test for Homogeneity of Variance (Operator A for C 117-49)	82
Bartlett Test for Homogeneity of Variance (Operator B for C 117-49)	84
Bartlett Test for Homogeneity of Variance (Operators A and B for C 117-49)	86
Calculations for Sum of Squares (Operator A for C 117-49)	87
Calculation of the Components of Variance for C 117-49	89
Tests of Hypotheses (Operator A for C 117-49)	90
Tests of Hypotheses (Operator B for C 117-49)	91
Bartlett Test for Homogeneity of Variance (Absorption for C 127-59)	92
Calculations for Sum of Squares (Absorption for C 127-59)	93
Calculation of Components of Variance (Absorption for C 127-59)	94
Tests of Hypotheses (Absorption for C 127-59)	95
Calculation of Control Limits for Between-Operators (Absorption for C 127-59)	96
Bartlett Test for Homogeneity of Variance (Absorption for C 128-57)	97
Calculations for Sum of Squares (Absorption for C 128-57)	98
Calculation of the Components of Variance (Absorption for C 128-57)	99
Tests of Hypotheses (Absorption for C 128-57)	100
Calculation of Control Limits for Between-Operators (Absorption for C 128-57)	101
Bartlett Test for Homogeneity of Variance (Bulk Specific Gravity for C 128-57)	102
Calculations for Sum of Squares (Bulk Specific Gravity for C 128-57)	103
Tests of Hypotheses (Bulk Specific Gravity for C 128-57)	104

TABLE OF CONTENTS (continued)

	Page
Bartlett Test for Homogeneity of Variance (Bulk SSD Specific Gravity for C 128-57)	105
Calculations for Sum of Squares (Bulk SSD Specific Gravity for C 128-57)	106
Bartlett Test for Homogeneity of Variance (Apparent Specific Gravity for C 128-57)	107
Calculations for Sum of Squares (Apparent Specific Gravity for C 128-57)	108

LIST OF TABLES

Table	Page
1. Gradation of Coarse Aggregate	32
2. Gradation of Fine Aggregate	32
3. Summary of Data for Test Method C 117-49	34
4. ANOV Table, Test Method C 117-49: Operator A	36
5. ANOV Table, Test Method C 117-49: Operator B	38
6. Means Determined for Data of Operator A	39
7. Summary of Data for Absorption, Test Method C 127-59	46
8. Summary of Data for Specific Gravity, Test Method 127-59	47
9. ANOV Table for Absorption Data, Test Method C 127-59	49
10. Sieve Analysis of Coarse Aggregate After Testing	55
11. Summary of Absorption Data, Test Method C 128-57	56
12. Summary of Specific Gravity Data, Test Method C 128-57 . . .	57
13. ANOV Table for Absorption Data, Test Method C 128-57	58
14. ANOV Table for Bulk Specific Gravity Data, Test Method C 128-57	65
15. ANOV Table for Bulk Saturated Surface Dry Specific Gravity Data, Test Method C 128-57	66
16. ANOV Table for Apparent Specific Gravity Data, Test Method C 128-57	67
17. Gradation of Fine Aggregate After Testing	69

LIST OF ILLUSTRATIONS

Figure	Page
1. Layout of the Experiment for Test Method C 117-49	22
2. Layout of the Experiment for Test Method C 127-59	23
3. Layout of the Experiment for Test Method C 128-57	24
4. Gradation of Fine Aggregate for Test Method C 128-57	30
5. Control Charts for Operator A, Test Method C 117-49	42
6. Control Charts for Operator B, Test Method C 117-49	44
7. Control Charts for Absorption Data of Limestone, Test Method C 127-59	51
8. Control Charts for Absorption Data of Gravel, Test Method C 127-59	52
9. Control Charts for Absorption Data of Fine Aggregate (F.M. = 2.40), Test Method C 128-57	61
10. Control Charts for Absorption Data of Fine Aggregate (F.M. = 3.00), Test Method C 128-57	62

ABSTRACT

Hanna, Steven John, M.S.C.E., Purdue University, June 1962. The Development of Precision Statements for Several ASTM Test Methods. Major Professor: J. F. McLaughlin.

Engineers have long been concerned with the precision of their so called "Standard" test methods. The inevitable question is, "within what limits can a person performing the Standard test on the same material expect his results to fall?" Many test methods have been devised but few of these give any indication of how well the results of repeated tests ought to agree. The problem of agreement between laboratories on the results from testing the same sample also arises.

Many standard methods of test for concrete and concrete aggregates published by the ASTM contain no information that will give the user of the test method a quantitative indication of how well individual test results of a series ought to agree with each other or how well test results from different laboratories, on supposedly "identical" samples, ought to agree. A statement that gives this information might be called a "precision statement". The objective of this work was to develop experimental designs with appropriate analyses from which these statements could be formulated.

To facilitate this objective several ASTM test methods were chosen for investigation (C 117-49, C 127-59, and C 128-57). For each of these test methods, a sample of aggregate was tested a number of times by two

operators acting independently. Prior to testing the variables associated with the test method were determined and a statistical model was written. After collecting the data the results were analyzed.

Using the results of the statistical analysis, confidence limits and control limits were determined at the 99 and 95 percent confidence levels. The precision statement was then formulated giving the plus and minus limits, the confidence levels at which these limits apply and the degrees of freedom associated with the limits.

During the repeated performance of the tests, observations were made leading to suggestions for modification of the procedure.

INTRODUCTION

Many standard methods of test for concrete and concrete aggregates (such as those published by the ASTM) contain little or no information that will give the user of the test method a quantitative indication of how well individual test results of a series ought to agree with each other or how well test results from different laboratories, on supposedly "identical" samples, ought to agree. Such a statement might be called a "precision statement". It has been suggested that a standard form of precision statement be included in ASTM Test Methods.

In a recent report to Subcommittee II-a (Evaluation of Data) of ASTM Committee C-9 (14)* it was suggested that ASTM designations be divided into four classes or groups. They are as follows:

- Group I - Methods of tests for which it appears possible to obtain a measure of repeatability and reproducibility having little or no sample variance either by repeating the test on the same sample or by making synthetic samples.
- Group II - Methods for which the measure of repeatability and reproducibility will necessarily include a component of variability introduced by sampling. That is, the method cannot be applied a number of times to the same sample

* Numbers in parenthesis refers to references listed in the bibliography.

and it does not appear that synthetic samples can be made.

Group III - Methods of test in which multiple specimens are required. These methods might possibly be classified with those of Group II but some differences should be recognized.

Group IV - Specifications and miscellaneous designations for which no statement of precision is needed.

In any test method there are many variables, some of which are inherent in the method itself and others that are the result of outside influence. Several of the components of variance that are present can be easily noted. There may be within-laboratory variance, variance due to differences in equipment, operators, etc. There may be between-laboratory variance and variance due to sample-to-sample differences. Finally there is the variability due to the method itself and that due to different materials.

As indicated by Group IV of the preceeding classification, not every ASTM designation requires precision statements of the type under consideration. Those relating to definition of terms do not. Others defining the for and generalized operational mode of equipment units do not. It is those "... methods-of-test specifications wherein apparatus and procedure is defined for evaluating a magnitude or a property that precision statements are needed. Such precision statements should define, within stated confidence limits, the maximum and the normal deviation between test results that may be expected of the test method in question when performed by experienced operators." (5). The ASTM Manual on Quality Control of Material (2) contains a section entitled "Presenting \pm Limits of Uncertainty

of an Observed Average". The necessary components of a precision statement can be determined using these two references as a basis. They seem to be: The \pm limits for a variable, the confidence level at which these limits apply, and the degrees of freedom associated with the determination of the limits. Since these limits are determined by statistical analysis of the data and sample statistics are used it is necessary to use the "Student's t" in determining confidence limits. This is mentioned by Kaplan (11) in his analysis and is generally discussed in detail in most statistical texts (i.e. 15).

Care should be taken when interpreting confidence limits on a sample statistic. A certain percent confidence limit means that if a number of test series are made and the sample statistic is computed and confidence limits are placed on the statistic, a certain percentage of the time these limits will include the true value. A common error is to say that a certain percent of the future sample statistics will be within the limits.

Using statistical analysis of test data, the degree of repeatability and reproducibility may be determined. Repeatability is defined as the quantitative measure of the variability associated with a single operator in a given laboratory (14). Reproducibility is defined as the quantitative measure of the variability associated with operators working in two different laboratories (14).

An experimental test is an attempt to estimate the true value of a given characteristic of a "population" of material for which there is a distribution of values for that characteristic. The population can be defined as all of the particular material under consideration. Samples taken from this population are generally dealt with in testing. The true

value of the characteristic is called a parameter. The estimate of this parameter obtained from a sample of the population is referred to as a sample statistic.

Usually in the presentation of data, a measure of central tendency and the variation from the true value are of primary concern. Perhaps one of the most common measures of central tendency is the mean, the arithmetic average of the values. The measure of variation or dispersion commonly used is the variance or mean square. It is defined as the sum of the squares of the deviates from the mean divided by one less than the total number of deviates. Sometimes the square root of the variance, the standard deviation, is used to describe the variation.

In many common statistical analyses it is assumed that the data being considered follow a normal distribution. Further, it is necessary that the samples be drawn from the population in a random fashion so that bias is not introduced.

A precision statement can be of great value to the user of the test. The limits provide a range of values in which the results of a duplicate test should lie with a certain probability. The operator is given a criterion for deciding if something has gone wrong. The variability of the test results, when compared with the expected variability, can give an indication of how well the test is being performed. That is, is the operator using the proper techniques and is he being accurate enough in his measurements as well as is he being careful enough?

If the repeatability and reproducibility of the test method is known, it is then possible to compare results of different operators and laboratories on a rational basis. Errors that are not within the inherent error

of the process or method can be discerned and with a knowledge that excessive errors are present the situation can be analyzed accordingly.

REVIEW OF LITERATURE

In recent years a great deal of interest has been generated in the problem of formulating a standard procedure for analysis of data in regards to the development of new test procedures or modification of old procedures. These analyses of data employ statistical methods in determining means, ranges, variances, etc. Several committees and subcommittees of the ASTM have investigated the problem and have recommended various procedures for analysis of data.

The analysis of data collected from both intralaboratory and interlaboratory studies has been deemed necessary before a new test procedure, or modification of an old procedure, may be accepted as a standard. The development of a new test procedure or a modification of an old procedure is initiated in a laboratory. From the development in one laboratory, where specific information must be determined, the investigation broadens to include several laboratories, as an interlaboratory study (18).

The attempt to find a standard method of attack has led to several papers on the subject.

Committee D-6 (Paper and Paper Products) has recommended a practice for interlaboratory evaluating of test methods used with paper and paper products (13). The following is a brief outline of the recommended procedure.

- A. Formulating study within one laboratory.
- B. Preliminary study within one laboratory.

- C. Study by task group.
- D. Pilot study.
- E. First interlaboratory.
- F. Main interlaboratory study.
- G. Decision of standardization.

Committee E-11 (on Quality Control of Materials) has also developed a recommended procedure for conducting an interlaboratory study of a test method (1). A discussion of the general requirements for conducting an interlaboratory study precedes a detailed discussion of experimental plans in the paper. It is recommended that a thorough examination of the test method be made and all available information about it be assembled and copies distributed. Particular attention is directed to the number and type of materials to be used in a study. It is emphasized that participants should have personnel who are familiar and experienced with the method under investigation. If they are not familiar with the method they should have an opportunity prior to the study to familiarize themselves with the method. It is further recommended that, for relatively new methods, or methods that have been greatly modified, a pilot study be made which will include all participants who will later take part in the main study.

The committee recommends that all test specimens should be prepared and allocated by one individual or group of individuals designated by the committee. This group would also devise data sheets for use during the study. The final analysis and interpretation should be left to a special group which includes persons familiar with basic methods of statistics. This group should analyze and interpret the data, then present it to the committee for discussion.

The recommendation includes detailed descriptions of interlaboratory studies using one, two, three, and many materials. Mathematical models are presented for the various plans and statistical analyses are also presented. Three sigma limits are to be determined as well as control limits.

A. B. Brown in a recent paper discusses the precision of present ASTM Tests on Bituminous paving materials (5). The paper presents a detailed discussion of the general status of precision statements of the ASTM specifications under the jurisdiction of Committee D-4 (Road and Paving Materials). Several forms of precision statements currently in use are presented and discussed. Dr. Brown recommends that a standard form of precision statement be developed for use by the ASTM and that a permanent record of the experimental data supporting such a statement for a test method be kept on file. Quoting from the paper:

"These features of the situation are, in themselves, bad, and the badness is compounded by the fact that there seems never to have been any policy regarding filing or publishing for future use the supporting data, however, inadequate, that evolved from the work of task forces and sub-committees doing the spade work in the development of these methods of test. The situation is still further confounded by the habit of revising specifications, other than editorially, from time to time. Usually these revisions at any one time are not drastic, and seldom are cooperative tests indulged in on any extensive scale to reconfirm earlier precision claims. Over a number of such revisions, it is quite possible that precision limits initially established should be changed."

A discussion of the statistical methods which may be used in analyzing test data is present and the distinction between repeatability and reproducibility is noted.

"In general, two categories of precision are commonly encountered, i.e., intra-operator and inter-laboratory precisions, respectively. The former deals with the variation between individual results performed by a given operator with a given apparatus, and the latter with the variation between the results of different operators working, usually at different sites, with different

apparatus. By common consent, the former type is generally referred to as the condition of repeatability and the latter type as the condition of reproducibility. Inasmuch as additional possibilities for variations rise for reproducibility over repeatability, the spread of values from the central value is ordinarily greater by several fold for reproducibility than for repeatability."

W. J. Youden in a recent paper discusses the evaluation of accuracy (19). He points out that good agreement among repeated measurements is no indication that the measure is close to the true value of the property being measured. There may be present a systematic error in the procedure which will tend to remove the measured average from the true average.

"One convenient viewpoint, whenever enough laboratories are involved, is to designate the average of all the laboratory averages as a grand average and the true value can be considered an estimate of the systematic error of the procedure."

The detection of systematic errors within one laboratory is not as easy as when the results of several laboratories are compared. Youden states that three major devices are commonly used to test a measurement procedure within one laboratory. They are:

1. Measurement of known materials.
2. Comparison with other measurement procedures.
3. Comparison with modifications of the given procedure.

M. E. Terry proposes that two theoretical approaches to the statistical treatment of research and development can be used together in the analysis of data (17). One method of analyzing

"... uses graphical methods wherein the data are first plotted in the pertinent recorded order in rational sub-groups, and the applicable control limits found from an average 'within-subgroup' estimate of dispersion. A subgroup central value and a dispersion estimate are plotted on charts together with their appropriate control limits."

The other approach

"... is to select a group of variables and a set of values of each variable, and then take measurements at selected combinations of these values. Then an estimate is made of the effect of changing each variable among its selected values, this effect being averaged over the selected values of each of the other variables. Randomization is used to average out the effects of the variables not under study."

W. J. Youden (18) presents a discussion of the problem beginning with the inception of the test procedure. A method of analysis using a graphical procedure which utilizes paired samples is discussed by Youden (20). Quoting from the summary:

A graphic procedure using paired samples has been developed to demonstrate to what extent systematic errors are present in interlaboratory tests. This procedure also indicates the extent of random errors and the ultimate precision of a test procedure.

A scoring system for evaluating the test results of a group of laboratories has been developed. It has been shown that the results of a few laboratories doing poor work can substantially inflate the apparent standard deviation obtained for any test.

J. R. Crandall and R. L. Blaine (9) in their study of interlaboratory cement tests utilized the Youden method.

The ASTM Manual on Quality Control (2) in a section entitled, "Presenting \pm Limits of Uncertainty of an Observed Average" gives procedures for determining \pm limits if the standard deviation (square root of the variance) is known. The "Student's t" is used to determine these limits.

I. W. Burr (6) in his text discussed the problem of Quality Control and presents procedures for determining if a process is in control. Control charts are used for this purpose with \pm limits placed on observed statistics.

The development of experimental designs for the formulation of test procedures using techniques of statistical analysis have become more and more common in recent years. Many text books have been published and many

papers written on the subject of statistical analysis. Increasing interest has been stimulated in obtaining good approximations for components of variance and other parameters. Of particular interest is the problem of confidence limits which may be placed on means, ranges, variances, etc. Analysis of variance is the primary statistical technique which has been used in the investigation of test procedures.

The March 1947, issue of Biometrics contains three papers which are of notable importance. The first by Churchill Eisenhart is entitled "The Assumptions Underlying the Analysis of Variance." The second by W. B. Cochran is entitled "Some Consequences when the Assumptions for Analysis of Variance are not Satisfied." The third by M. S. Bartlett is entitled "The Use of Transformations."

Of particular interest is the article by Cochran in which he lists the assumptions required for the analysis of variance and discusses the consequences when these assumptions are not met. Quoting from the summary and conclusions:

"The analysis of variance depends on the assumptions that the treatment and environmental effects are additive and that the experimental errors are independent in the probability sense, have equal variance and are normally distributed. Failure of any assumption will impair to some extent the standard properties on which the widespread utility of the technique depends."

Quoting further:

"In general, the factors that are liable to cause the most severe disturbances are extreme skewness, the presence of gross errors, anomalous behavior of certain treatments or parts of the experiment, marked departures from the additive relationship, and changes in the error variance, either related to the mean or to certain treatments of parts of the experiment."

The choice of proper design of the experiment is a primary prerequisite for a statistical analysis which will give the desired information.

Texts such as those by Steel and Torrie (15) and Ostle (12) discuss a number of design methods. Detailed procedures are given for the calculation of components of variance and for tests of hypotheses.

Ostle (12) gives the following advantages and disadvantages in his discussion of factorial design models:

Advantages:

1. Greater efficiency in the use of available experimental resources is achieved.
2. Information is obtained about the various interactions.
3. The experimental results are applicable over a wider range of conditions; that is, due to the combining of the various factors in one experiment the results are of a more comprehensive nature.
4. There is a gain due to the hidden replication arising from the factorial arrangement.

Disadvantages:

1. The experimental setup and the resulting statistical analysis are more complex.
2. With a large number of treatment combinations the selection of homogeneous experimental units becomes more difficult.
3. Certain of the treatment combinations may be of little or no interest; consequently, some of the experimental resources may be wasted.

It seems to be the general opinion that the advantages far outweigh the disadvantages. Quoting from Ostle, p. 372 (12) "Although we have listed almost as many disadvantages as advantages for factorials, it is

felt that the advantages far outweigh the disadvantages in importance. Factorials are therefore highly recommended as useful devices in scientific experimentation in all fields."

PURPOSE AND SCOPE

The purpose of this investigation was to develop experimental designs by which data could be collected to formulate precision statements for certain ASTM Test Methods and to collect these data for the Methods under consideration. Statistical analyses of the data were made and from these, confidence limits and control limits were determined at the 95 and 99 percent confidence levels. The investigation was intended to determine the within-laboratory precision of the test methods for different operators following exactly the procedures for the methods as given in the ASTM Standards. During the testing a careful analysis of the testing procedure was made to determine if the method is adequate and if the desired results are achieved by the procedure.

Since precision statements (as they have been defined) contain confidence intervals and degrees of freedom it was necessary to evaluate the means and variances of the test results. From these, limits on the sample statistics were determined for the two confidence levels.

Three ASTM Test Methods falling in Group I (as previously defined) were used.

They were:

1. C 117-49 Test for Amount of Material Finer than No. 200 Sieve in Aggregate
2. C 127-59 Test for Specific Gravity and Absorption of Coarse Aggregate
3. C 128-57 Test for Specific Gravity and Absorption of Fine Aggregates.

STATISTICAL PROCEDURES

To accomplish the stated objective there are, generally speaking, many experimental designs that could be used. The choice of the model must be made by the experimenter but the choice will necessarily be influenced by the information he is seeking. The experimenter must select the model that will give him the maximum information and at the same time be both flexible and economical. The simpler the model the better. A complicated model is a waste of time and effort if the same results can be achieved using a simpler model. Many times an experimenter has collected data, taken it to a statistician for analysis and, much to his dismay, found that the data were useless in trying to test his hypothesis. This can be eliminated by a careful study of the problem before testing has begun.

In keeping with this line of thought, each test method was examined to determine variables present and choose a statistical model for the analysis. The model for each method was determined to give the desired information about the test method being investigated. For the three methods investigated, the use of factorial models seemed to be advantageous. Ostle (12) in his discussion of factorial models states the following advantages and disadvantages:

Advantages:

1. Greater efficiency in the use of available experimental resources is achieved.
2. Information is obtained about the various interactions.

3. The experimental results are applicable over a wider range of conditions; that is, due to the combining of the various factors in one experiment the results are of a more comprehensive nature.
4. There is a gain due to the hidden replication arising from the factorial arrangement.

Disadvantages:

1. The experimental setup and the resulting statistical analysis are more complex.
2. With a large number of treatment combinations the selection of homogeneous experimental units becomes more difficult.
3. Certain of the treatment combinations may be of little or no interest; consequently, some of the experimental resources may be wasted.

It seems to be the general opinion that the advantages far outweigh the disadvantages. Quoting from Ostle, p. 372 (12) "Although we have listed almost as many disadvantages as advantages for factorials, it is felt that the advantages far outweigh the disadvantages in importance. Factorials are therefore highly recommended as useful devices in scientific experimentation in all fields."

In this investigation sampling variation was eliminated by performing the test on the same sample a number of times. Since different operators would be using the test method, operators were included as a variable, a random variable, since the results would be generalized to include all operators. Type of material was also included as a variable and here again results would be generalized to include all materials to be tested by the method and hence it would be a random variable.

GENERAL ANALYSIS OF VARIANCE FOR A 3-FACTOR FACTORIAL IN A RANDOMIZED

COMPLETE BLOCK DESIGN: COMPONENT-OF-VARIANCE MODEL

Source of Variation	Degree of Freedom	Mean Square	Expected Mean Square
Replicates	$r-1$	$R_{yy}/(r-1)$	$\sigma^2 + abc \sum_{i=1}^r p_i^2/(r-1)$
Treatments			
A	$a-1$	$A_{yy}/(a-1)$	$\sigma^2 + r\sigma_{\alpha\gamma}^2 + rc\sigma_{\alpha\theta}^2 + rb\sigma_{\alpha\gamma}^2 + rbc\sigma_{\alpha}^2$
B	$b-1$	$B_{yy}/(b-1)$	$\sigma^2 + r\sigma_{\alpha\theta}^2 + rc\sigma_{\alpha\theta}^2 + ra\sigma_{\theta\gamma}^2 + rac\sigma_{\theta}^2$
C	$c-1$	$C_{yy}/(c-1)$	$\sigma^2 + r\sigma_{\alpha\gamma}^2 + rb\sigma_{\alpha\gamma}^2 + ra\sigma_{\theta\gamma}^2 + rab\sigma_{\gamma}^2$
AB	$(a-1)(b-1)$	$(AB)_{yy}/(a-1)(b-1)$	$\sigma^2 + r\sigma_{\alpha\theta\gamma}^2 + rc\sigma_{\alpha\theta}^2$
AC	$(a-1)(c-1)$	$(AC)_{yy}/(a-1)(c-1)$	$\sigma^2 + r\sigma_{\alpha\theta\gamma}^2 + rb\sigma_{\alpha\gamma}^2$
BC	$(b-1)(c-1)$	$(BC)_{yy}/(b-1)(c-1)$	$\sigma^2 + r\sigma_{\alpha\theta\gamma}^2 + ra\sigma_{\theta\gamma}^2$
ABC	$(a-1)(b-1)(c-1)$	$(ABC)_{yy}/(a-1)(b-1)(c-1)$	$\sigma^2 + r\sigma_{\alpha\theta\gamma}^2$
Experimental error	$(r-1)(abc-1)$	$E_{yy}/(r-1)(abc-1)$	σ^2
Total	$abc - 1$		

(After Ostle)

The two previously mentioned variables were present in all three test methods investigated. In addition to these two variables the amount of minus 200 material was included as a variable for test method C 117-49. The range from 1% to 10% was chosen as being of primary importance and four levels were selected within this range. This variable was determined to be a fixed variable since the extremes were represented and two other levels covered the range sufficiently.

The mathematical models for the three test methods are presented on the following pages. In addition, Figure 1 shows the layout of the experiment for test method C 117-49, Figure 2 shows the layout of the experiment for test method C 127-59, and Figure 3 shows the layout of the experiment for test method C 128-57.

Prior to testing, an Analysis of Variance Table (ANOV) was formulated for each test method and the expected mean squares (EMS) were determined.

A Type I error, usually denoted by α , is committed if the hypothesis is rejected when it is actually true. Obviously one wishes to choose a test in which the probability of an error of this kind is small. It has been found generally acceptable in practice to use an α -level of 0.05, although this depends on the type of problem under consideration. For this study an α -level of 0.05 has been used in performing F-tests.

The following equation was used in calculating the variance used in computing the control limits.

$$\text{Variance } (\bar{Y}_{jkl}) = \frac{\sigma_{\alpha}^2}{n_{\alpha}} + \frac{\sigma_{\beta}^2}{n_{\beta}} + \frac{\sigma_{\alpha\beta}^2}{n_{\alpha\beta}} + \frac{\sigma_{\rho}^2}{n_{\rho}} + \frac{\sigma_{\epsilon}^2}{n_{\epsilon}}$$

where:

σ_{α}^2 = Variance component due to α

σ_{β}^2 = Variance component due to β

σ_{ρ}^2 = Variance component due to ρ

$\sigma_{\alpha\rho}^2$ = Variance component due to the interaction of α and ρ

σ_{ϵ}^2 = Variance component due to experimental error

and,

n_{α} = number of a's = a

n_{ρ} = number of b's = b

$n_{\alpha\rho}$ = a times b

n_{ρ} = number of replicates = r

n_{ϵ} = total number of observations used to estimate variance (\bar{Y}_{jkl}), in this case abr

The Variance (\bar{Y}_{jkl}) is used in computing control limits which are placed on the "grand" mean. If control limits are placed on means which are averaged over only one of the variables then the appropriate variance (\bar{Y}_x) must be used. For example, if control limits are placed on means averaged over replicates only the variance would be Variance (\bar{Y}_{ρ}) = $\frac{\sigma_{\rho}^2}{n_{\rho}} + \frac{\sigma_{\epsilon}^2}{n_{\epsilon}}$. Where n_{ϵ} is equal to the number of observations used to estimate the variance, in this case $n_{\epsilon} = n_{\rho} = r$.

Mathematical Model for C 117-49

$$Y_{ijkl} = \mu + \rho_i + \alpha_j + \beta_k + \gamma_l + (\alpha\beta)_{jk} + (\alpha\gamma)_{jl} + (\beta\gamma)_{kl} + (\alpha\beta\gamma)_{jkl} + \epsilon_{ijkl}$$

Y_{ijkl} = an individual obs. in the i th replicate, in the j th level of factor a, in the k th level of factor b, and in the l th level of factor c

μ = the true mean value

ρ_i = effect of the i th replicate

α_j = effect of the j th level of factor a (material)

β_k = effect of the k th level of factor b (amount)

γ_l = effect of the l th level of factor c (operators)

$(\alpha\gamma)_{jl}$ = effect of the interaction of the j th level of factor a with the l th level of factor c

$(\alpha\beta)_{jk}$ = effect of the interaction of the j th level of factor a with the k th level of factor b

$(\beta\gamma)_{kl}$ = effect of the interaction of the k th level of factor b with the l th level of factor c

$(\alpha\beta\gamma)_{jkl}$ = effect of interaction of the three levels of factors

ϵ_{ijkl} = effect of the experimental unit in the i th replicate to which the (jkl) th treatment combination has been randomly assigned.

and the terms α_j , γ_l , $(\alpha\gamma)_{jl}$ and ϵ_{ijkl} are assumed to be independently normally distributed with expected values of 0 and variances σ_α^2 , σ_γ^2 , $\sigma_{\alpha\gamma}^2$, and σ^2 respectively,

$$\text{and } \sum_{i=1}^r \rho_i = \sum_{k=1}^4 \beta_k = \sum_{k=1}^4 (\alpha\beta)_{jk} = \sum_{k=1}^4 (\beta\gamma)_{kl} = \sum_{k=1}^4 (\alpha\beta\gamma)_{jkl} = 0$$

Mathematical Model for C 127-59 and C 128-57

$$Y_{ijk} = \mu + \rho_i + \alpha_j + \beta_k + (\alpha\beta)_{jk} + \epsilon_{ijk}$$

Y_{ijk} = an individual observation in the i th replicate, in the j th level of factor a and in the k th level of factor b

μ = mean effect

ρ_i = effect of the i th replicate

α_j = effect of the j th level of factor a (material)

β_k = effect of the k th level of factor b (operator)

$(\alpha\beta)_{jk}$ = effect of the interaction of the j th level of factor a with the k th level of factor b

ϵ_{ijk} = effect of the experimental unit in the i th replicate to which the (jk) th treatment combination has been randomly assigned

The various terms of the model (except μ and ρ_i) are all assumed to be independently and normally distributed with expected values of 0 and variances σ_α^2 , σ_ρ^2 , $\sigma_{\alpha\beta}^2$, and σ^2 , respectively.



3-FACTOR FACTORIAL DESIGN

REPLICATE	OPERATOR	LIMESTONE				CLAY			
		1%	4%	7%	10%	1%	4%	7%	10%
1	A								
	B								
.	A								
	B								
.	A								
	B								
10	A								
	B								

$i = 1, 2, \dots, 10$ (Replicates) Material - Random
 $j = 1, 2$ (Material) Amount - Fixed
 $k = 1, 2, 3, 4$ (Amount) Operators - Random
 $l = 1, 2$ (Operators)

Fig. 1. Layout of the Experiment for Test Method C 117-49.

2-FACTOR FACTORIAL DESIGN

REPLICATE	LIMESTONE		GRAVEL	
	OPERATOR A	OPERATOR B	OPERATOR A	OPERATOR B
1				
.				
.				
10				

$i = 1, 2, \dots, 10$ (Replicates)

$j = 1, 2$ (Material)

$k = 1, 2$ (Operators)

Material - Random

Operators - Random

Fig. 2. Layout of the Experiment for Test Method C 127-59.

2-FACTOR FACTORIAL DESIGN

REPLICATE	F.A. (FM = 2.40)		F.A. (FM = 3.00)	
	OPERATOR A	OPERATOR B	OPERATOR A	OPERATOR B
1				
.				
.				
10				

$i = 1, 2, \dots, 10$ (Replicates)

$j = 1, 2$ (Material)

$k = 1, 2$ (Operators)

Material - Random

Operators - Random

Fig. 3. Layout of the Experiment for Test Method C 128-57.

PROCEDURES

The procedures detailed in the ASTM Standards were followed exactly by the operators when performing the tests. Two operators, acting independently, performed ten replicates for each test method investigated. The operators performed several practice tests to familiarize themselves with the method and test procedures before actual testing began.

From the data collected by the operators, sample means and sample variances were calculated. A statistical analysis was made on the data for each test method. Confidence limits were determined for 95 and 99 percent confidence levels. From the statistical analysis of the data, precision statements of various kinds could be formulated for each test method studied.

Test Method C 117-49

The purpose of C 117-49 is to determine the amount of material finer than the No. 200 sieve size in aggregate. Minimum sample sizes of 500 grams, 2500 grams, and 5000 grams are required for aggregate whose nominal maximum sizes are No. 4(4760-micron), 3/4 inch, and 1 1/2 inches or over, respectively. Weight determinations are to be made to the nearest 0.02 percent. Percent of material finer than No. 200 sieve size is calculated based on the original dry weight of the sample.

The procedure outlined in the test method is to first dry the sample to a constant weight*. Next the sample is weighed then placed in a container large enough to hold it plus the wash water when subjected to vigorous

* Test Method C 117-49 specifies $230 \pm 90^{\circ}\text{F}$

agitation. Water is added, the mixture is agitated and the wash water is then poured over a No. 200 sieve. This procedure is repeated until the wash water is clear. All material which has been retained on the No. 200 sieve is returned to the container which is then placed in an oven and the sample is dried to a constant weight. The final dry weight is determined and calculation of percent finer than No. 200 sieve is made.

The material used in the performance of this test was local concrete sand of glacial origin. The sand was prepared by washing it over a No. 200 sieve several times to eliminate all material finer than that sieve size. The sand was dried to a constant weight and material finer than the No. 200 sieve was added to batches of it in amounts so as to produce 500.0 gram samples containing one, four, seven, and ten percent material finer than the No. 200 sieve.

Two types of minus 200 material were used in the investigation. In one series limestone dust, obtained from commercial limestone filler, was used. In the other series, a limestone-residual clay soil was the source of the fine material*. The limestone dust was chosen so as to represent a non-cohesive fine material and the clay was chosen so as to represent a very cohesive fine material.

The proper amount of minus 200 material was "weighed-out" on a triple beam balance to obtain an accuracy of 0.1 gram and added to the samples of sand. After addition of the minus 200 material, each sample was mixed thoroughly and then tested in accordance with the procedures outlined in C 117-49. All weight determinations were made to the nearest 0.1 gram.

For each replicate, eight samples were prepared and tested in random order. Each replicate contained samples representing the two material

* Unified Soil Classification, CH.



types and the four levels of each material. A table of random numbers was used to achieve the random order. This precaution was taken to eliminate any bias which might occur due to an ordered procedure of preparing and testing the samples. Replicates were made on different days to prevent any undue influence from the operator's "mood" on any one particular day from influencing the final results.

Test Method C 127-59

The purpose of C 127-59 is to determine the twenty-four hour absorption, bulk specific gravity, bulk specific gravity (saturated surface dry basis) and the apparent specific gravity of coarse aggregate. Coarse aggregate is defined as that portion of an aggregate retained on the No. 4 sieve (3). A sample of approximately 5 kg. of aggregate is required. The sample is washed to remove dust and other coatings from the surface of the particles, dried to a constant weight, and then immersed in water for a period of twenty-four hours. After this period, the sample is removed from the water, rolled in a large absorbent cloth until a saturated surface dry condition is achieved, immediately weighed, then immersed in water and weighed. After the immersed weight is obtained, the sample is dried to a constant weight, cooled and weighed. Weight determinations are made to the nearest one-half gram.

Two materials were used in this experiment. One was a crushed limestone from a quarry near Bloomington, Indiana, the other was a local gravel. The coarse aggregates were graded so as to have equal amounts of material retained on the 1, 3/4, 1/2, 3/8 inch, and No. 4 sieves (Table 1). They were sieved, washed and dried to a constant weight, then the various fractions were combined to form a 5000.0 gram sample.

After the samples of coarse aggregate were prepared, they were tested for absorption and specific gravity following the procedures detailed in C 127-59. All weight determinations were made to the nearest one-half gram.

For each replicate the two samples were tested by each operator. Replicate determinations were made on different days to prevent bias. Upon completion of the testing the two samples were again sieved and any change in weight of each fraction was noted.

Test Method C 128-57

The purpose of C 128-57 is to determine the twenty-four hour absorption, bulk specific gravity, bulk specific gravity (saturated surface dry basis) and the apparent specific gravity of fine aggregate. Fine aggregate is defined as that portion of an aggregate passing the No. 4 sieve and retained on the No. 200 sieve (3). A sample of approximately 1000.0 grams is required. The sample is dried to a constant weight, covered with water and allowed to stand for twenty-four hours. At the end of this period the sample is air dried to a saturated surface dry condition and a sample of 500.0 grams is placed in a 500 ml. volumetric flask. Enough water is added to bring the level almost to the 500 ml. mark. The flask is rolled to remove entrapped air and is then placed in a constant temperature bath. At the end of one hour the flask is removed and filled to the 500 ml. mark and the weight (or volume) of water added is determined. The sample is removed from the flask placed in a container and dried to a constant weight. The dry weight is determined and calculation of absorption and specific gravities are made.

The material used for this was local concrete sand. The sand was washed to remove material finer than the No. 200 sieve, dried to a constant weight, and separated into various fractions using a logarithmic series of sieves. Sand from these fractions was then recombined to form samples having two gradations, one with a fineness modulus of 2.40 and the other a fineness modulus of 3.00 (Table 2). Each sample weighed 1000.0 grams. The two gradations are shown in Figure 4. The limits for fine aggregate as given in Standard Specifications for Concrete Aggregates (C 33-57) are also noted in Figure 4.

Each of the two samples was then tested for absorption and specific gravity following the procedures outlined in test method C 128-57. One procedure not specifically noted in the ASTM Standards was used, however. When the sample was removed from the 500 ml. flask the water was removed by decanting the sample. The sand plus the filter paper was then placed in the drying oven together and the dry weight obtained. The filter paper was oven dried prior to testing to a constant weight and its dry weight obtained. By subtracting this weight from the combined dry weight the sample dry weight was obtained. This procedure was used to speed drying time and while the method does not explicitly state that the procedure should be followed, the wording does not preclude its use.

The portion of the sand not used in the flask was placed in a separate container and dried to a constant weight along with the sample from the flask. After the dry weights were determined the portions were recombined for the next replicate. As before, two operators acting independently performed ten replicates on each sample. One hundred percent ethanol was used in minor amounts to reduce foaming in the flask.

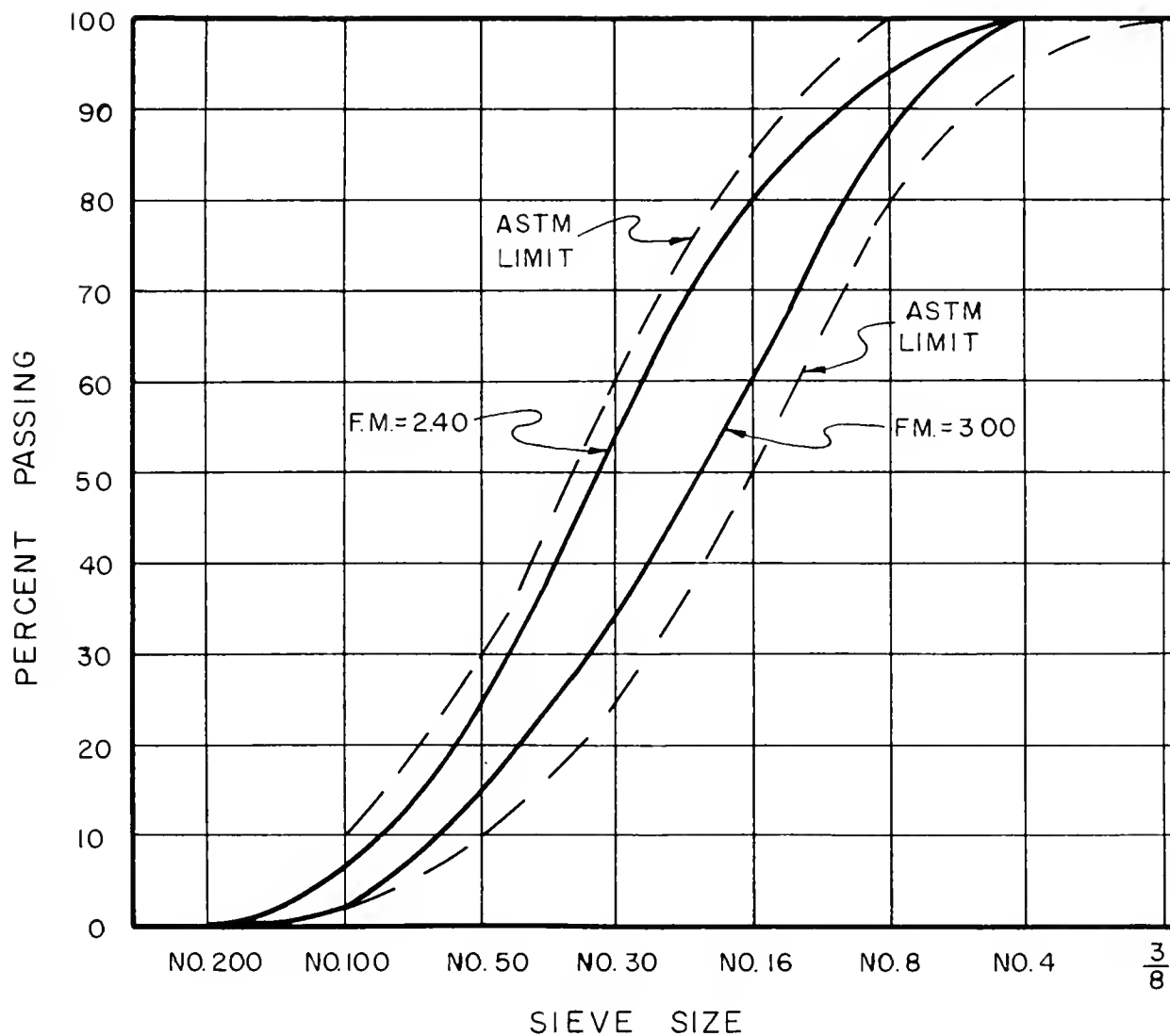


FIG. 4 GRADATION OF FINE AGGREGATE FOR
TEST METHOD C 128-57.

Upon completion of testing the two samples were sieved and the change in weight of each fraction noted. Loss in weight occurred during handling of the sample.

TABLE 1

Gradation of Coarse Aggregates

Sieve Fraction	Limestone	Gravel
1 1/2 in. - 1 in.	1000.0 gm.	1000.0 gm.
1 in. - 3/4 in.	1000.0 gm.	1000.0 gm.
3/4 in. - 1/2 in.	1000.0 gm.	1000.0 gm.
1/2 in. - 3/8 in.	1000.0 gm.	1000.0 gm.
3/8 in. - #4	<u>1000.0 gm.</u>	<u>1000.0 gm.</u>
	5000.0 gm.	5000.0 gm.

TABLE 2

Gradation of Fine Aggregates

Sieve Fraction	Sample #1 (F.M.=2.40)	Sample #2 (F.M.=3.00)
3/8 in. - #4	0	0
#4 - #8	60.0 gm.	120.0 gm.
#8 - #16	140.0 gm.	280.0 gm.
#16 - #30	250.0 gm.	250.0 gm.
#30 - #50	300.0 gm.	200.0 gm.
#50 - #100	190.0 gm.	130.0 gm.
#100 - #200	<u>60.0 gm.</u>	<u>20.0 gm.</u>
	1000.0 gm.	1000.0 gm.

DISCUSSION OF RESULTS

The results of this study and the discussion of these results have been divided into four parts. The first part is concerned with the analysis of the data for Test Method C 117-49. The second part is concerned with the analysis of the data for Test Method C 127-59; the third part with the analysis of data for Test Method C 128-57, and the fourth part with the discussion of all test results and their application to the development of precision statements. The second and third parts have been further divided into two parts, one concerned with the analysis of absorption data, and the other concerned with the analysis of specific gravity data.

Results of Test Method C 117-49

It should be recalled that a 3-factor factorial model was chosen for test method C 117-49. Operators and material were random variables and amounts of minus 200 material was a fixed variable. A complete tabulation of the data collected for this method is shown in Table 3. The layout of this Table is the same as that shown in Figure 1 for the layout of the experiment.

A Bartlett test for homogeneity of variance was performed on the data for Operator A and resulted in accepting the hypothesis that the within-sample variances were equal. That is, the eight variances for two materials and four levels of each were accepted as being equal at an α -level of 0.05. The same conclusion was reached for the within-sample

TABLE 3
SUMMARY OF DATA FOR TEST METHOD C 117-49

Replicate	Operator	Limestone ^{a1}				Clay ^{a2}			
		1% _{b1}	4% _{b2}	7% _{b3}	10% _{b4}	1% _{b1}	4% _{b2}	7% _{b3}	10% _{b4}
1	A c ₁	0.98	4.04	6.90	10.14	1.26	3.94	6.80	9.88
	B c ₂	1.06	4.02	7.02	10.02	1.00	4.00	7.00	10.02
2	A	1.04	4.30	6.98	10.02	1.00	4.06	7.12	10.04
	B	0.96	3.94	6.98	9.88	1.04	3.90	7.00	10.00
3	A	1.14	4.20	7.26	10.20	1.26	3.42	6.88	10.08
	B	1.02	3.96	6.98	9.90	0.98	4.04	6.98	10.06
4	A	1.00	3.96	6.84	9.94	0.96	3.84	6.90	9.84
	B	1.06	3.92	6.90	10.00	1.00	4.00	7.06	10.00
5	A	1.00	4.00	6.94	9.96	1.02	3.84	6.92	9.94
	B	1.00	4.04	7.16	10.04	0.92	3.92	6.88	10.00
6	A	1.14	4.06	6.94	10.14	1.06	4.00	7.10	10.06
	B	1.02	3.94	7.00	10.08	1.08	4.00	7.02	9.90
7	A	1.12	3.92	7.10	9.92	1.06	4.08	6.98	10.00
	B	1.06	3.96	6.92	10.00	0.98	4.06	6.96	10.06
8	A	1.00	4.08	7.00	10.02	1.04	3.98	6.90	9.96
	B	1.00	3.92	6.90	9.90	0.94	4.00	6.94	9.96
9	A	1.14	4.00	7.10	9.98	1.04	3.88	6.98	10.08
	B	1.04	4.00	7.08	10.00	1.04	4.00	7.00	9.98
10	A	1.10	3.96	7.00	9.98	1.10	4.08	6.98	10.00
	B	1.02	4.04	7.00	10.02	1.06	4.04	7.06	10.02

variances of Operator B. However, a Bartlett test over the entire table of data (including both operators) resulted in the rejection of the hypothesis that all within-sample variances were equal at the 0.05 α -level.

With the problem of non-homogeneous variance present, it was deemed advisable to analyze the results of Operator A and Operator B separately since one of the assumptions underlying the analysis of variance is that variances are homogeneous (10). Analyzing the data for Operator A and Operator B separately eliminated the operator variable and hence changed the statistical model from a 3-factor factorial to two 2-factor factorial models similar to the model for analysis of test method C 127-59.

The calculations for the sum of squares for the analysis of variance tables followed the procedure presented in Ostle (12). These calculations may be noted in detail in Appendix A. Only the sums of squares themselves are presented in the analysis of variance tables. The mean squares are found by dividing the sums of squares by their degrees of freedom and since the expected mean squares may be determined by method given in most statistics texts the individual components of variance may be determined. See Appendix A for these calculations.

Table 4 presents the analysis of Variance Table for Operator A, Test Method C 117-49. The estimate of experimental error, σ^2 , is shown to be 0.0124. This mean square is used in testing hypotheses in regard to the effect of replicates, materials, and the operator-material interaction. The mean square for amounts was, as would be expected, very large, 296.1773, in comparison to the variance which was used to test the hypothesis that $\sigma_{\epsilon}^2 = 0$ (amount effect). This effect must be tested using the mean square for amount-material interaction since the expected mean square for amount includes σ^2 and $\sigma_{\alpha\epsilon}^2$.

TABLE 4

ANALYSIS OF VARIANCE TABLE

TEST METHOD C 117-49: OPERATOR A

Source of Variation	Degrees of Freedom	Sum of Squares	Mean Square	Expected Mean Square
Replicates	$(r-1) = 9$	$R_{yy} = 0.1819$	0.0202	$\sigma^2 + ab \sigma_e^2$
α	$(a-1) = 1$	$A_{yy} = 0.0594$	0.0594	$\sigma^2 + rb \sigma_\alpha^2$
β	$(b-1) = 3$	$B_{yy} = 888.5318$	296.1773	$\sigma^2 + r \sigma_{\alpha\beta}^2 + ra \sum_{k=1}^{K=4} \frac{\sigma_k^2}{(b-1)}$
$\alpha\beta$	$(a-1)(b-1) = 3$	$(AB)_{yy} = 0.0609$	0.0203	$\sigma^2 + r \sigma_{\alpha\beta}^2$
Experimental Error	$(r-1)(ab-1) = 63$	$E_{yy} = 0.7851$	0.0124	σ^2
Total	$(rab-1) = 79$			

F-tests performed on the data from the analysis of variance indicated that materials and amounts were significant (see Appendix A for calculations). That is that we could not, at the 0.05 α -level accept the hypothesis that $\sigma_{\alpha}^2 = 0$ (material effect) and that $\sigma_{\beta}^2 = 0$ (amount effect). The results of the F-tests indicated further that we could accept the hypothesis that $\sigma_r^2 = 0$ (replicate effect) and $\sigma_{\alpha\beta}^2 = 0$ (material-amount interaction).

Table 5 presents a similar analysis of Variance Table Operator B, Test Method C 117-49. The estimate of experimental error was 0.0028. Since the expected mean squares for this design model are the same as those for the previously discussed model the same procedures of hypothesis testing were followed. As before the mean square for amounts were very large, 298.8232, with respect to the mean square used in testing the hypothesis that $\sigma_{\beta}^2 = 0$, namely $\sigma^2 + \sigma_{\alpha\beta}^2$.

F-tests performed on the data from the analysis of variance of the data for Operator B indicated that amounts were significant, $\sigma_{\beta}^2 \neq 0$, at 0.05 α -level.

Amounts would be expected to be significant since their means were chosen to be different. The fact that the material effect was significant for Operator A indicated that the type of minus 200 material has some effect on the results obtained. An examination of the means for Operator A indicates that generally a higher percentage of minus 200 material was determined for the finely ground limestone than for minus 200 clay material (Table 6).

TABLE 5

ANALYSIS OF VARIANCE TABLE

TEST METHOD C 117-49: OPERATOR B

Source of Variation	Degrees of Freedom	Sum of Squares	Mean Square	Expected Mean Square
Replicates	$(r-1) = 9$	$R_{yy} = 0.0490$	0.0054	$\sigma^2 + ab \sigma_\rho^2$
α	$(a-1) = 1$	$A_{yy} = 0.0003$	0.0003	$\sigma^2 + rb \sigma_\alpha^2$
β	$(b-1) = 3$	$B_{yy} = 896.4696$	298.8232	$\sigma^2 + r \sigma_{\alpha\beta}^2 + ra \sum_{i=1}^{i=10} \frac{\beta_k^2}{(b-1)}$
$\alpha\beta$	$(a-1)(b-1) = 3$	$(AB)_{yy} = 0.0075$	0.0025	$\sigma^2 + r \sigma_{\alpha\beta}^2$
Experimental Error	$(r-1)(ab-1) = 63$	$E_{yy} = 0.1738$	0.0028	σ^2
Total	$(rab-1) = 79$			

TABLE 6
MEANS DETERMINED FOR DATA OF OPERATOR A

Material	Limestone				Clay			
Percent Added	1	4	7	10	1	4	7	10
Percent Determined by C 117-49	1.066	4.052	7.006	10.030	1.080	3.912	6.952	9.988

Since material was found not significant for the data of Operator B, coupled with the fact that the variances of Operator A and Operator B were found to be different, it is possible that the results obtained by the test method are highly sensitive to the skill of the operator performing the test. Another factor, lends some support to this suggestion. Operator B was an "experienced" operator while Operator A was not (even though he was given a reasonable "break-in" or learning period). Therefore, the results obtained by Operator B may be indicative of results obtained when the test method is performed by an experienced laboratory technician, while the results obtained by Operator A may be indicative of results obtained when the test method is performed by a person not trained in laboratory techniques of this nature.

The problem of establishing an acceptable range within which results of duplicate determination should agree, or establishing an interval, within which one expects, with some given probability, the population parameter being estimated to lie, was complicated by the non-homogeneity of variance encountered.

Another factor that arises is selecting the type of limits to be set, i.e. confidence limits or control limits. Confidence limits on a mean or

observation provide a range that, with a given probability, include the true mean, μ . Control limits on the other hand provide limits in which a certain percent of the sample statistics should fall in the long run. A set of data that has a number of observations outside the control limits (greater than that amount determined by choice of an α -level) is said to be out of control. Several factors may account for the data being out of control. One can be the fact that some assignable cause had led to discrepancies between observations, another is that the precision of the test method or process is incapable of producing results of the quality required.

For the purpose of this study it seems appropriate to present both confidence limits and control limits. These limits are presented for the data Operator A and Operator B separately.

Confidence Limits for Operator A

$s^2 = 0.0124$ from Table 4. Therefore $s = 0.111$.

Ninety-five percent confidence limits on the mean of n observations made to determine the amount of material finer than No. 200 sieve are $\pm t_{.05} \frac{s}{\sqrt{n}}$, where t is based on the chosen α -level (0.05 in this case) and the degrees of freedom associated with s . In this case the number of degrees of freedom is 63. For these conditions, the value of t is 1.998 and the limits are $\pm \frac{0.222}{\sqrt{n}}$.

If, for example, three observations are made the confidence limits placed on the average of the three is $\bar{Y} \pm \frac{0.222}{\sqrt{3}} = \bar{Y} \pm 0.128$. The interpretation that is placed on these limits is that one is 95% confident that the true mean, μ , lies within the limits established by, $\bar{Y} \pm 0.128$ where \bar{Y} is the observed mean.

Ninety-nine percent confidence limits are $\pm t_{.01} \frac{s}{\sqrt{n}}$ which for the determined s and 63 degrees of freedom is $\pm \frac{0.295}{\sqrt{n}}$. The same interpretation as previously noted above may be made.

Control Limits for Operator A

The establishment of control limits for a measurement process can give a basis for determining whether observations are made within the limits of variability allowable or inherent in the process. The control limits were placed on individual observations. Since both material and amount were determined significant, the limits were placed about the average of ten replicates. The limits, placed on individual observations are wider than limits placed on means. The limits are indicative of how well an operator may repeat his measurements.

The limits are as follows:

$$\text{Estimate of variance} = \sigma_{Y_{jr}}^2 = \frac{\sigma_e^2}{1} + \frac{\sigma^2}{1} = 0.0134$$

For 0.05 α -level based on n observation, where n equals ten replicates, $t_{.05} = 2.262$. The control limits are then equal to:

$$\pm 2.262 \sqrt{0.0134} = \pm 0.262\%$$

For 0.01 α -level, $t = 3.25$ and the control limits are:

$$\pm 3.25 \sqrt{0.0134} = \pm 0.376\%$$

A plot of the observations in each replicate and the control limits as determined above are shown in Figure 5 for the clay material. A similar chart can be constructed for the limestone material except that the mean, \bar{X} , would be shifted.

Confidence Limits for Operator B

$s^2 = 0.0028$ from Table 5. Therefore, $s = 0.0548$.

95% confidence limits equal: $\pm \frac{0.110}{\sqrt{n}}$ following the same procedures

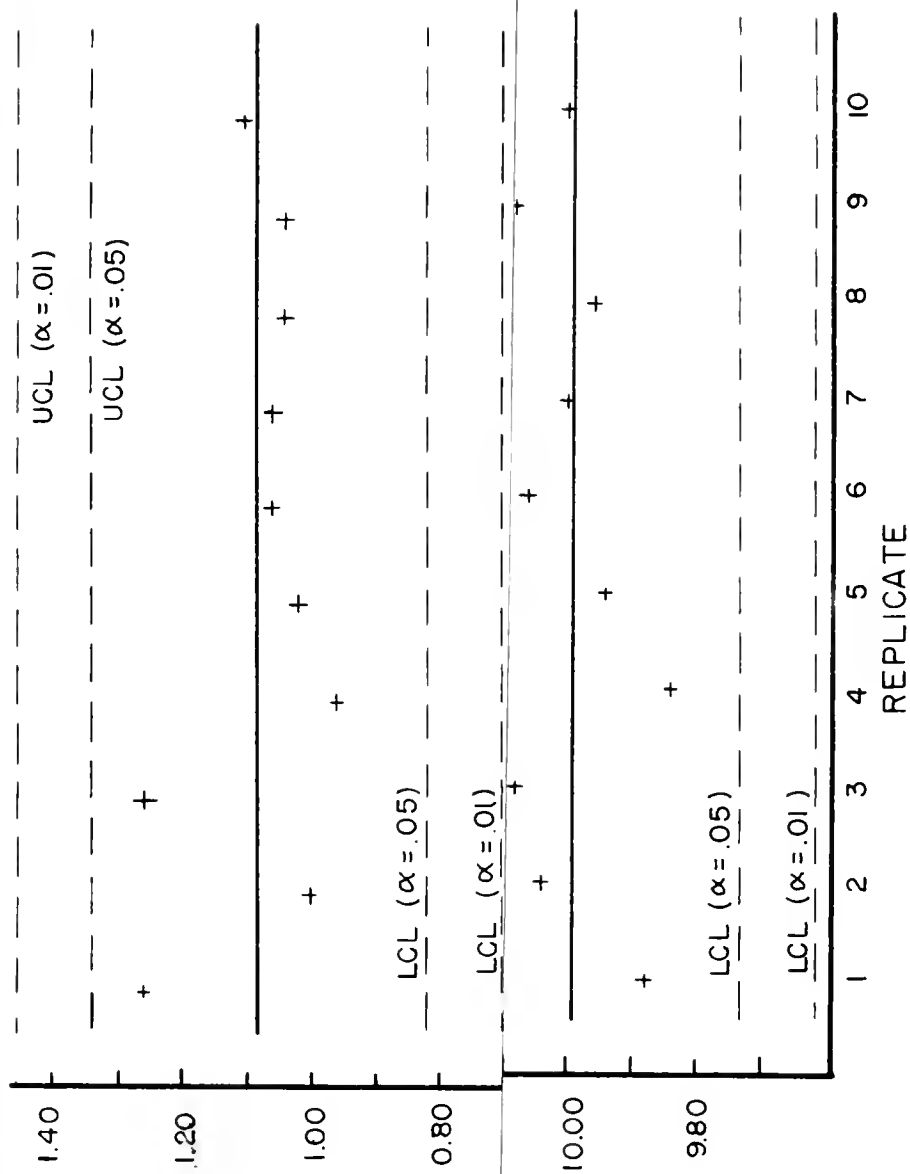


FIG. 5 CONTROL CHARTS FOR OPERATOR A, TEST METHOD C 117-49.

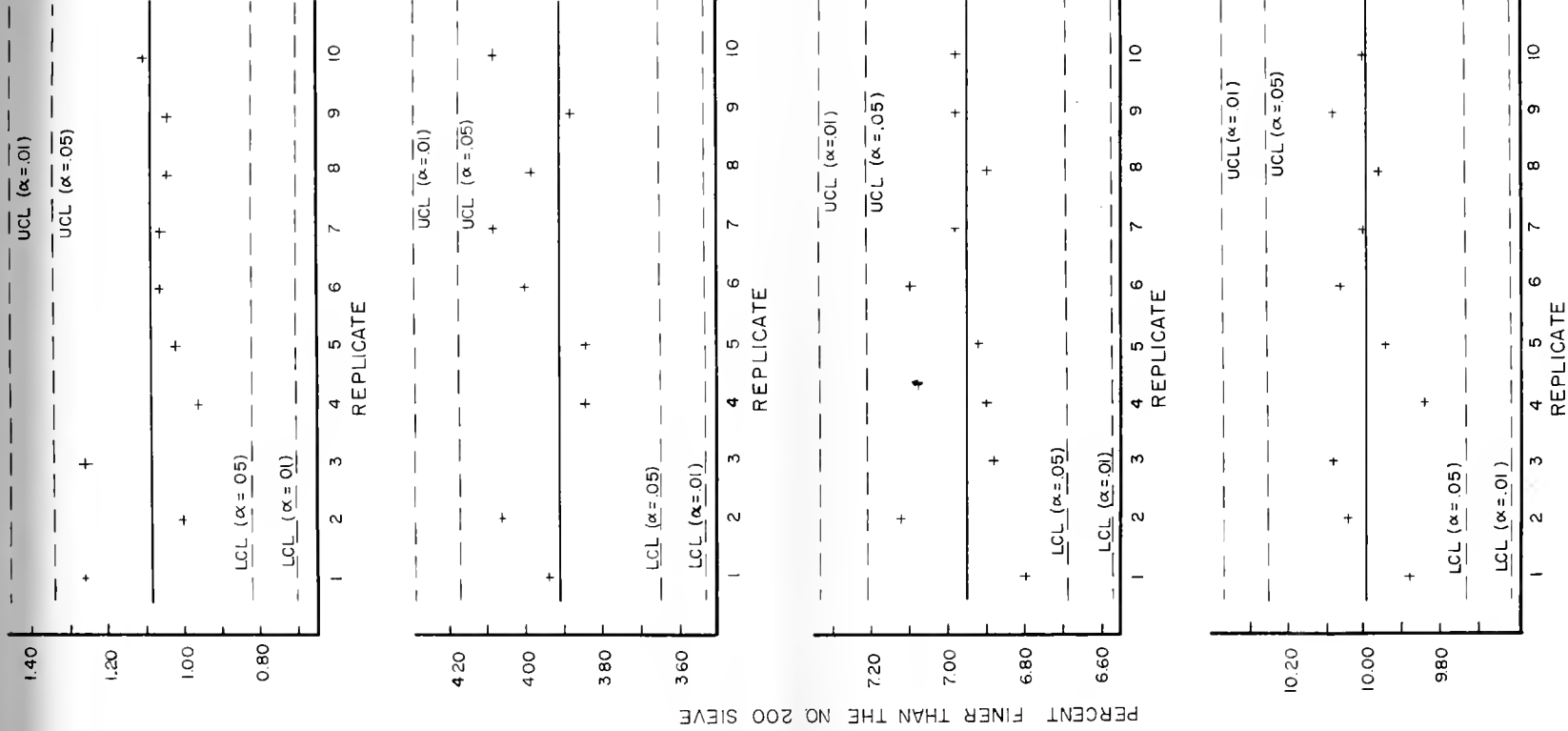


FIG. 5 CONTROL CHARTS FOR OPERATOR A, TEST METHOD C 117-49.

shown in the determination of confidence limits for Operator A. Similarly, 99% confidence limits equal: $\pm \frac{0.145}{\sqrt{n}}$

Control Limits for Operator B

Control limits for data of Operator B were placed on individual observations. The limits were plotted about means averaged over replicates and materials since amounts were significant. The procedures followed in determining these limits were the same as those used for determining control limits for the data of Operator A except that, obviously a different variance was used (see Appendix A). The limits are:

$\pm 0.116\%$ for an α -level of 0.05

$\pm 0.159\%$ for an α -level of 0.01

Figure 6 shows the observations and the control limits as determined above.

General Observations

Figure 5 and Figure 6 give an indication of how wide the limits are and how some of the data were scattered. This scattering could be due to several factors (e.g. a variable which was not controlled, such as laboratory humidity, may be significant). It is of interest to note that the ASTM Committee responsible for this test method has proposed a revision*. The proposed revision requires that weight determination and calculation of percent finer than No. 200 sieve be made to the nearest 0.1 percent instead of the nearest 0.02 percent now designated. The remainder of the proposed revision appears to be in keeping with procedures followed during this study. If the percent finer than No. 200 sieve were determined to the nearest 0.1 percent, as suggested, the variances are, for practical purposes, zero for the data of Operator B. The situation is then similar to

* Accepted as a tentative, C 117-61T, in 1961.

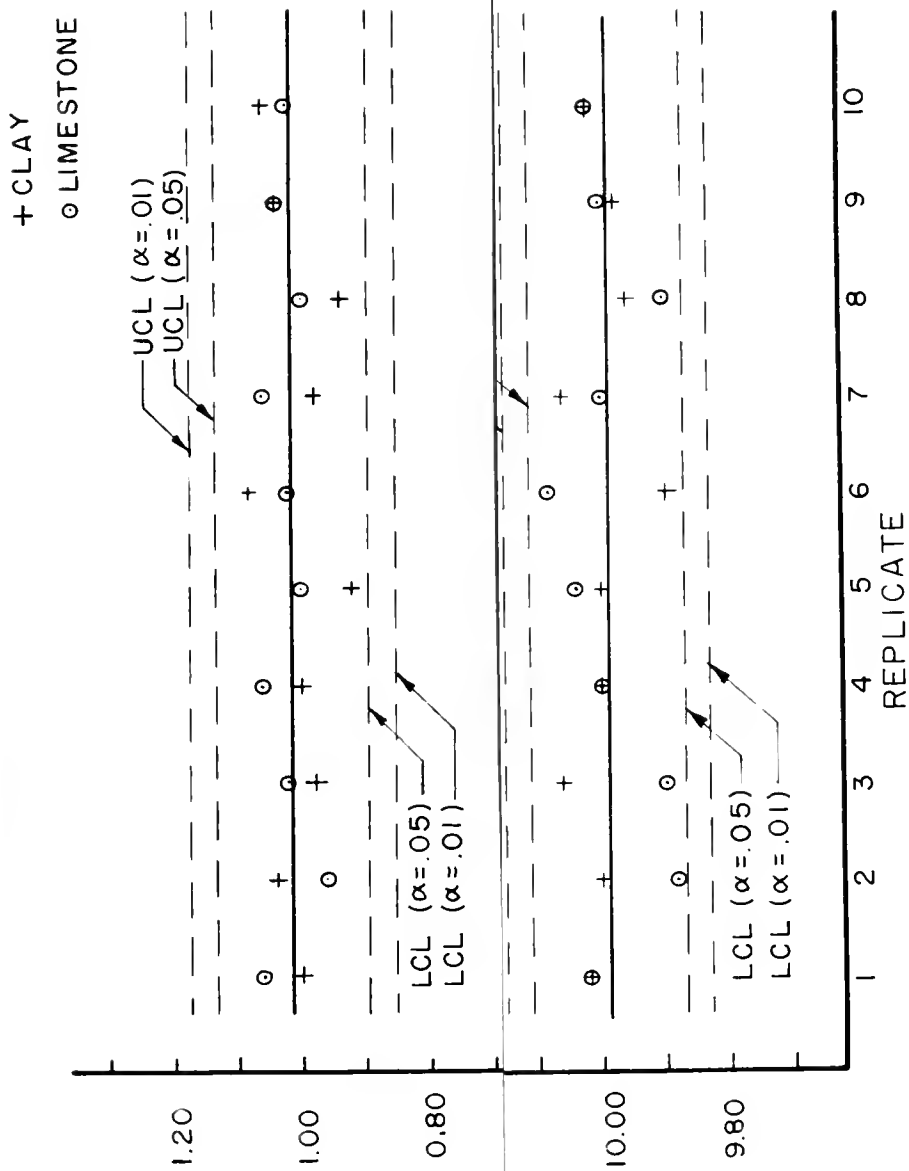


FIG. 6 CONTROL CHARTS FOR OPERATOR B, TEST METHOD C 117-49

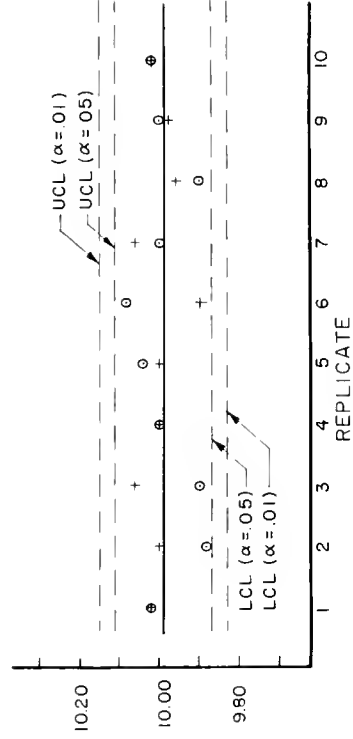
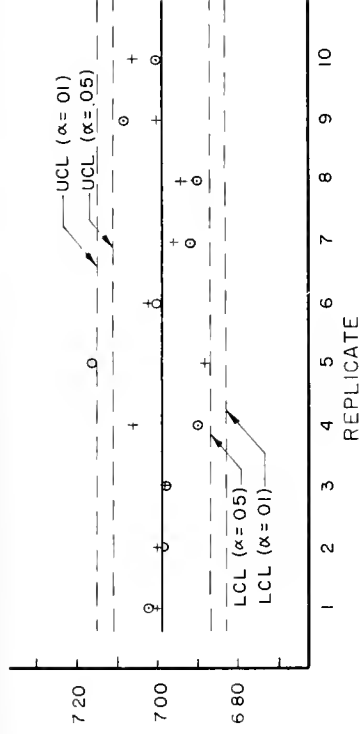
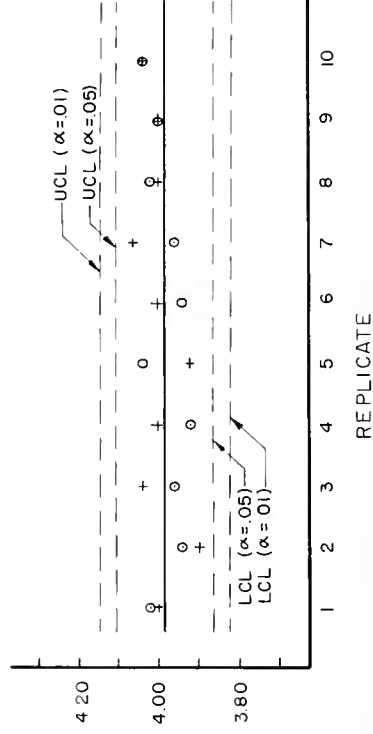
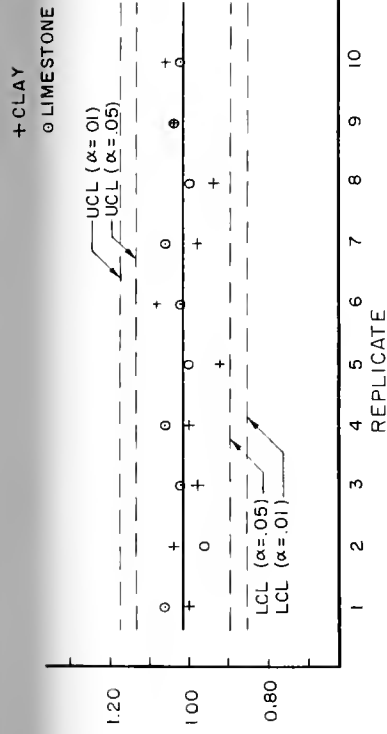


FIG. 6 CONTROL CHARTS FOR OPERATOR B, TEST METHOD C 117-49

the situation encountered for the specific gravity data of Test Method C 127-59. Namely, the data are insensitive to statistical analyses and practical limits of ± 0.1 percent seem appropriate. Noting what happened for the data of Operator A the placement of these "practical" limits may not be justified.

During the course of this study several interesting factors were noted. In the preparation of the minus 200 material a sieve which showed some visible signs of wear was used. When the test method itself was performed on the samples containing this minus 200 material, new sieves were used. Results showed that the material removed was consistently less than that which was added. As a result of this, the minus 200 material was checked using the new sieves and a considerable amount was found to be retained on the new sieves, all of which had passed the older sieve. Hence one may conclude that the condition of the sieves is an important factor in determining accurately the quantity of minus 200 material present in a sample and it is important that sieves which meet the Standards specified by the Test Method be used.

It was also noted that it required more washings per sample if the sample was simply stirred. However, if a combination of stirring and hydraulic pressure was used, a more efficient washing was achieved.

Results of Test Method C 127-59

Prior to testing, the variables present were determined and a 2-factor factorial model was chosen for the analysis of absorption and specific gravity data. A tabulation of the absorption data with the column means indicated is given in Table 7. A tabulation for the specific gravity data is given in Table 8.

TABLE 7
SUMMARY OF DATA FOR ABSORPTION
TEST METHOD C 127-59

Replicate	(LS) Non-Absorptive Aggregate		(GR) Absorptive Aggregate	
	Operator A	Operator B	Operator A	Operator B
1	0.90	0.94	1.22	1.25
2	0.87	0.91	1.24	1.20
3	0.87	0.92	1.34	1.25
4	0.85	0.93	1.31	1.26
5	0.89	0.87	1.34	1.22
6	0.87	0.90	1.25	1.18
7	0.85	0.87	1.36	1.26
8	0.92	0.89	1.36	1.26
9	0.94	0.83	1.34	1.27
10	0.92	0.89	1.32	1.28
\bar{x}	0.888	0.895	1.308	1.243

TABLE 8
SUMMARY OF DATA FOR SPECIFIC GRAVITY
TEST METHOD C 127-59

Mat.	Operator A			Operator B		
	Bulk	Bulk (SSD)	Apparent	Bulk	Bulk (SSD)	Apparent
GR	2.64	2.67	2.73	2.64	2.67	2.73
	2.63	2.66	2.72	2.64	2.67	2.73
	2.62	2.66	2.73	2.64	2.67	2.73
	2.63	2.67	2.73	2.64	2.67	2.73
	2.63	2.67	2.73	2.64	2.67	2.73
	2.64	2.67	2.73	2.64	2.68	2.72
	2.63	2.67	2.73	2.64	2.68	2.73
	2.63	2.67	2.73	2.64	2.68	2.74
	2.63	2.67	2.73	2.64	2.67	2.73
	2.64	2.67	2.73	2.64	2.67	2.73
LS	2.64	2.66	2.71	2.64	2.67	2.71
	2.65	2.67	2.71	2.65	2.67	2.71
	2.65	2.67	2.71	2.65	2.67	2.71
	2.65	2.67	2.71	2.65	2.67	2.71
	2.65	2.68	2.70	2.65	2.67	2.71
	2.65	2.67	2.71	2.65	2.67	2.71
	2.65	2.67	2.71	2.65	2.67	2.71
	2.65	2.67	2.71	2.65	2.68	2.72
	2.64	2.67	2.71	2.66	2.68	2.72
	2.65	2.67	2.71	2.65	2.67	2.71

Absorption

A Bartlett test for homogeneity of variance was performed on the data noted in Table 7. The test resulted in accepting the hypothesis that within sample variances were equal at the 0.05 α -level. With these variances determined equal, it was then possible to proceed with the analysis of the data using the factorial model.

The sums of squares and the mean squares were calculated according to the procedures previously given for 2-factor factorial models. These sums of squares and mean squares are shown in Table 9. The calculations for the sums of squares are given in detail in Appendix A. As before the expected mean squares are also given in Table 9. The estimate of experimental error is 0.00146 as determined by the ANOV.

F-tests performed on the data from the analysis of variance indicated that there was a significant interaction between materials and operators, i.e. $\sigma_{\alpha\theta}^2 \neq 0$. Replicates and operators were not significant. The F-tests also indicated that materials were not significant, $\sigma_{\alpha}^2 = 0$. Materials were expected to be significant since aggregates having different absorptions were deliberately selected. The significant interaction can be logically explained. One operator may be more experienced, or more familiar, with one material than the other and this could conceivably influence his results.

With estimates of the variance components available, limits were determined. As before, both confidence limits and control limits were determined. Control limits were placed on individual observations. This provides an estimate of the precision which an individual operator performing the test would be expected to have.

TABLE 9

ANOV TABLE FOR ABSORPTION DATA

TEST METHOD C 127-59

Source of Variation	Degrees of Freedom	Sum of Squares	Mean Square	Expected Mean Square
Replicates	$(r-1) = 9$	$R_{yy} = 0.0128$	$\frac{R_{yy}}{9} = 0.00142$	$\sigma^2 + ab\sigma_\epsilon^2$
α	$(a-1) = 1$	$A_{yy} = 1.4745$	$\frac{A_{yy}}{1} = 1.47450$	$\sigma^2 + r\sigma_{\alpha\beta}^2 + rb\sigma_\alpha^2$
β	$(b-1) = 1$	$B_{yy} = 0.0084$	$\frac{B_{yy}}{1} = 0.00840$	$\sigma^2 + r\sigma_{\alpha\beta}^2 + ra\sigma_\beta^2$
$\alpha\beta$	$(a-1)(b-1) = 1$	$(AB)_{yy} = 0.0130$	$\frac{(AB)_{yy}}{1} = 0.01300$	$\sigma^2 + r\sigma_{\alpha\beta}^2$
Experimental Error	$(r-1)(ab-1) = 27$	$E_{yy} = 0.0394$	$\frac{E_{yy}}{27} = 0.00146$	σ^2
Total	$(rab-1) = 39$			

The calculation of the components of variance results in $\sigma_e^2 = 0$. The sum of squares for replicates and experimental error were pooled to form a better estimate of the population variance. This variance was determined to be 0.00145 with 36 degrees of freedom. Therefore, control limits placed on the individual observations were determined to be, $\pm t \sqrt{\frac{0.00145}{1}}$. The value of t was determined by the α -level chosen and $(n-1)$ degrees of freedom. The value of n for this case was equal to the number of replicates, 10. The control limits for $\alpha = 0.05$ and $\alpha = 0.01$ are indicated below.

For $\alpha = 0.05$.

$$t_{0.05, 9 \text{ degrees of freedom}} = 2.262$$

$$\text{control limits} = \pm 2.262 \sqrt{\frac{0.00145}{1}} = \pm 0.086\%$$

For $\alpha = 0.01$.

$$t_{0.01, 9 \text{ degrees of freedom}} = 3.250$$

$$\text{control limits} = \pm 3.250 \sqrt{\frac{0.00145}{1}} = \pm 0.133\%$$

Figure 7 and 8 show these limits placed about the means of ten replicates.

Confidence limits were determined as follows. $\sigma^2 = 0.00145$ (pooled), therefore $\sigma = 0.038$. Ninety-five percent confidence limits are $\pm t \frac{\sigma}{\sqrt{n}}$, where t is based on the chosen α -level and the degrees of freedom associated with σ . For this case, t equals 2.029 based on 0.05 α -level and 36 degrees of freedom. The confidence limits are therefore, $\pm 2.029 \frac{0.038}{\sqrt{n}} = \pm \frac{0.077}{\sqrt{n}}$. The interpretation of this is that we are 95% confident that the true mean, μ , lies within the limits estimated by $\bar{Y} \pm \frac{0.077}{\sqrt{n}}$, where \bar{Y} is the observed mean based on n observations.

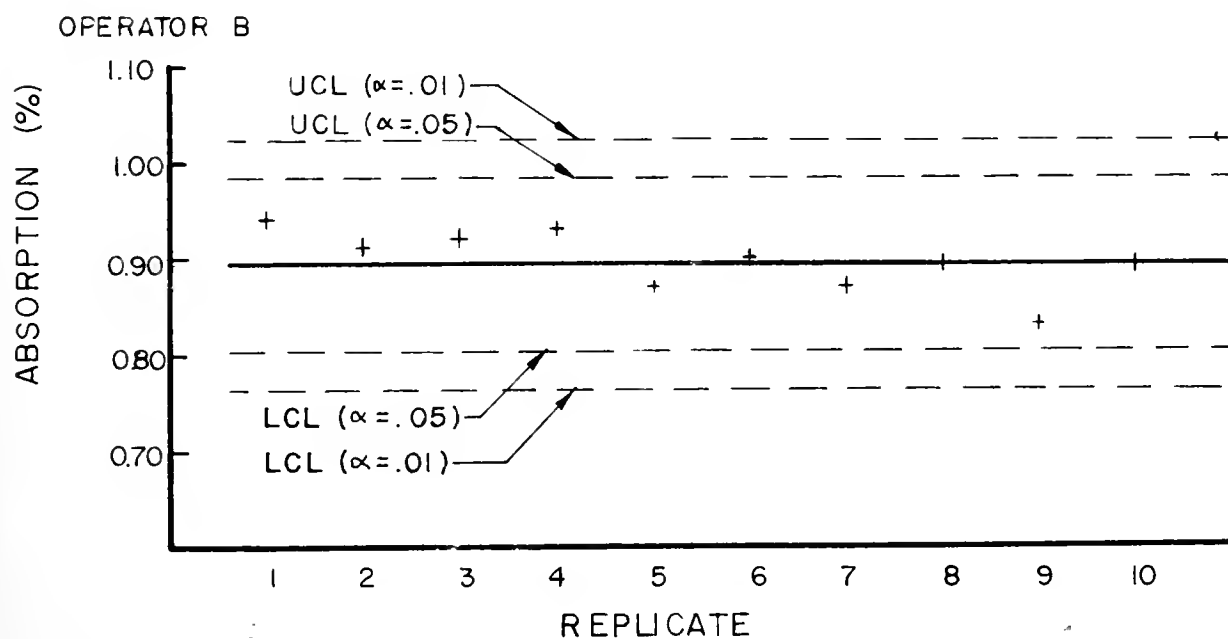
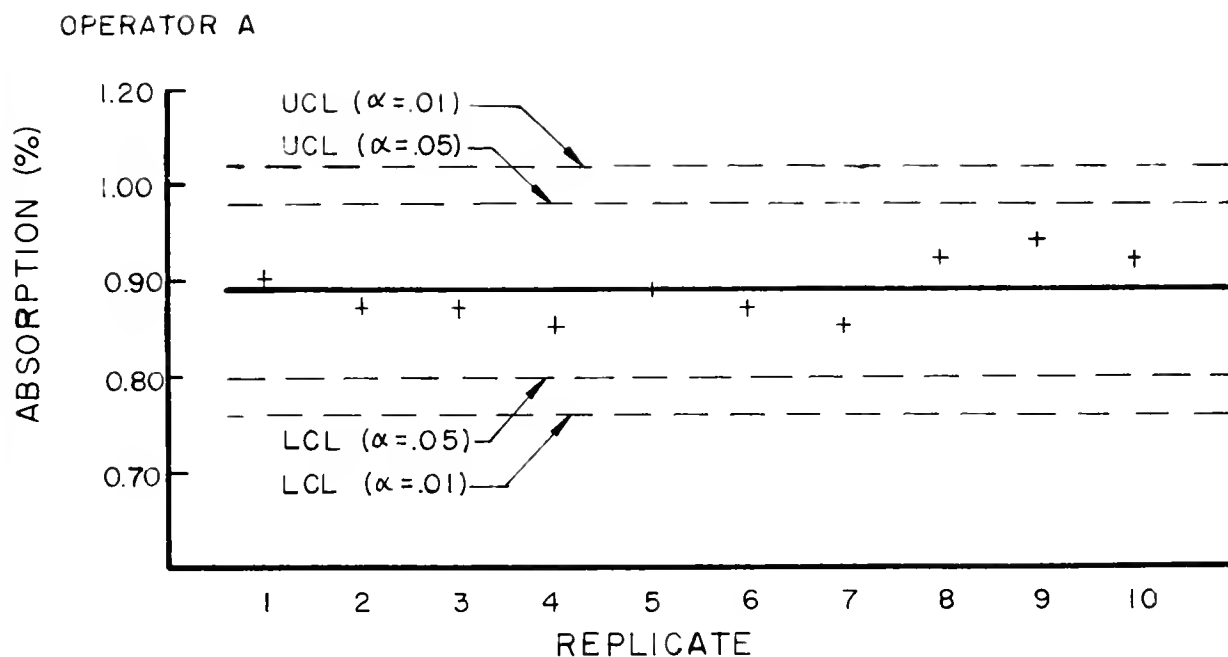


FIG. 7 CONTROL CHARTS FOR ABSORPTION DATA OF LIMESTONE, TEST METHOD C 127-59

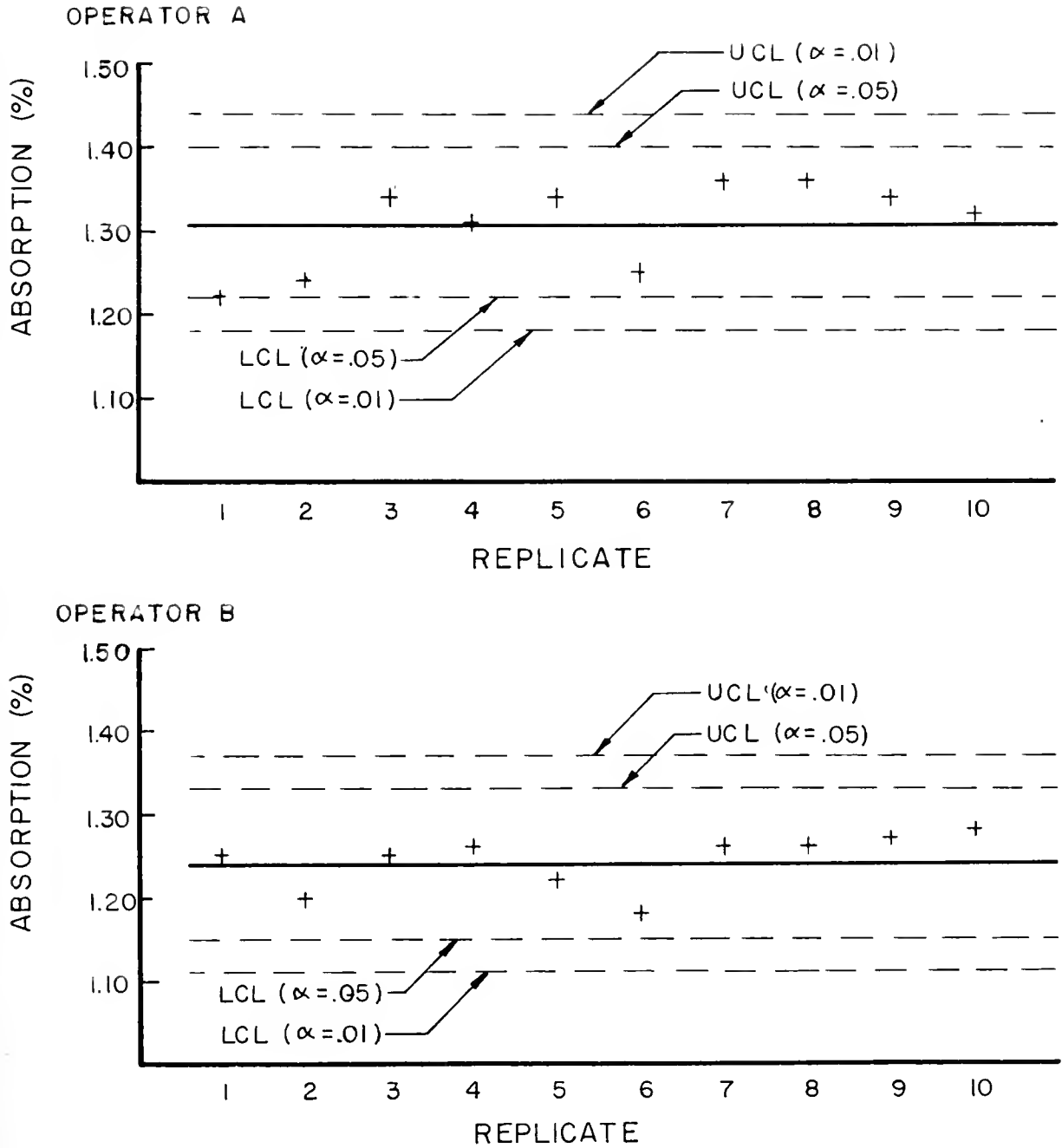


FIG. 8 CONTROL CHARTS FOR ABSORPTION DATA OF GRAVEL, TEST METHOD C 127-59.

Ninety-nine percent confidence limits are $\pm t_{.01} \frac{\sigma}{\sqrt{n}}$. For an α -level of 0.01 and 36 degrees of freedom, $t = 2.722$. The confidence limits are $\pm 2.722 \frac{0.038}{\sqrt{n}} = \pm \frac{0.104}{\sqrt{n}}$. The interpretation of this is, that we are 99% confident that the true mean, μ , lies within the limits estimated by $\bar{Y} \pm \frac{0.104}{\sqrt{n}}$, where \bar{Y} is the observed mean based on n observations.

Control limits were also determined for "between operators". These limits were determined as follows.

The estimate of variance, $\sigma_{\bar{Y}_j}^2$, equals 0.000073. The limits are $\pm t \sqrt{0.000073}$. For an α -level of 0.05 and 27 degrees of freedom, t equals 2.052 and the limits are $\pm 0.018\%$. For an α -level of 0.01 and 27 degrees of freedom, t equals 2.771 and the limits are $\pm 0.024\%$.

These limits are narrower than the control limits determined for a single operator and are applicable to situations where one is interested in making comparisons between operator means.

Specific Gravity

A statistical analysis of the specific gravity data (Table 8) was not warranted since the range in values for each type of specific gravity never exceeded 0.02 and the variances were, for practical purposes, zero. The practical control limits for this case, therefore, are ± 0.01 . These limits would also apply to between-operators since a check of the data indicated almost identical specific gravity determinations by each operator.

A precision statement as it might appear in the test method would be as follows: Duplicate determinations made by the same, or by different operators, should lie within the range determined by $\bar{Y} \pm 0.01$, where \bar{Y} is the observed average of the specific gravity determinations.

General Observations

One particular point was noted during the course of the study of Test Method C 127-59. This was that care should be taken, when submersing the sample into the water for the weight determination in water, to place the sample in the water in such a manner as to prevent as much as possible the entrapment of air. Possibly some reference in the test method is warranted.

The sieve analysis of the coarse aggregate, after testing, is given in Table 10. The total loss in weight was, generally, evenly distributed over the testing period.

Results of Test Method C 128-57

Prior to testing, the variables present were determined and a 2-factor factorial model was chosen for the analysis of absorption and specific gravity data. A tabulation of the absorption data with the column means indicated is given in Table 11. A tabulation for the specific gravity data is given in Table 12.

Absorption

A Bartlett test for homogeneity of variance was performed on the data noted in Table 11. The test resulted in accepting the hypothesis that within-sample variances were equal at the 0.05 α -level. With these variances determined equal, it was then possible to proceed with the analysis of the data using the factorial model.

The sums of squares and the mean squares were calculated according to the procedures previously given for 2-factor factorial models. These sums of squares and mean squares are shown in Table 13. The calculation for the sums of squares are given in detail in Appendix A. As before the

TABLE 10
SIEVE ANALYSIS OF COARSE AGGREGATE AFTER TESTING

	LS	Gravel
1 1/2 - 1	1000.0 gm	1000.0 gm
1 - 3/4	970.5 gm	1000.0 gm
3/4 - 1/2	941.0 gm	947.5 gm
1/2 - 3/8	945.0 gm	952.0 gm
3/8 - #4	1076.5 gm	1066.5 gm
Pass #4	<u>24.5 gm</u>	<u>1.0 gm</u>
	4957.5 gm	4966.0 gm

Original dry wt., 5000.0 gm for both samples

Note: See summary of data for indication of how loss in wt. was distributed over the testing period.

TABLE 11
SUMMARY OF ABSORPTION DATA
TEST METHOD C 128-57

Replicate	Absorption, in %			
	Operator A		Operator B	
	FM 3.00	FM 2.40	FM 3.00	FM 2.40
1	1.24	1.01	1.19	0.81
2	1.19	1.09	1.42	0.87
3	1.17	0.99	1.38	1.21
4	1.24	0.93	1.34	1.30
5	1.17	1.05	1.32	1.32
6	1.03	0.85	1.36	1.28
7	1.17	0.91	1.26	1.20
8	0.85	0.97	1.40	1.17
9	0.83	0.73	1.46	1.40
10	1.17	1.07	1.36	1.13
\bar{X}	1.106	0.960	1.349	1.169

TABLE 12
SUMMARY OF SPECIFIC GRAVITY DATA
TEST METHOD C 128-57

FM	Operator A			Operator B		
	Specific Gravities			Specific Gravities		
	B	BSSD	A	B	BSSD	A
3.00	2.60	2.63	2.68	2.60	2.63	2.68
	2.58	2.62	2.67	2.58	2.62	2.68
	2.59	2.62	2.67	2.59	2.62	2.68
	2.60	2.64	2.69	2.59	2.63	2.69
	2.60	2.63	2.68	2.58	2.61	2.67
	2.62	2.65	2.70	2.59	2.63	2.69
	2.60	2.63	2.68	2.60	2.63	2.69
	2.62	2.64	2.68	2.59	2.63	2.69
	2.62	2.64	2.68	2.58	2.61	2.68
	2.60	2.63	2.68	2.58	2.62	2.68
2.40	2.60	2.63	2.67	2.62	2.64	2.67
	2.61	2.64	2.69	2.60	2.62	2.66
	2.62	2.65	2.69	2.59	2.63	2.68
	2.61	2.63	2.67	2.60	2.64	2.69
	2.60	2.63	2.67	2.59	2.63	2.68
	2.63	2.65	2.69	2.59	2.63	2.68
	2.62	2.64	2.68	2.60	2.63	2.68
	2.61	2.64	2.68	2.60	2.63	2.68
	2.62	2.64	2.67	2.59	2.63	2.69
	2.63	2.65	2.70	2.60	2.63	2.68

TABLE 13
ANALYSIS OF VARIANCE TABLE FOR ABSORPTION DATA
TEST METHOD C 128-57

Source of Variation	Degrees of Freedom	Sum of Squares	Mean Square	Expected Mean Square
Replicates	$(r-1) = 9$	$R_{yy} = 0.0896$	0.00995	$\sigma^2 + ab \sigma_e^2$
α	$(a-1) = 1$	$A_{yy} = 0.2657$	0.2657	$\sigma^2 + r \sigma_{\alpha\beta}^2 + rb \sigma_\alpha^2$
β	$(b-1) = 1$	$B_{yy} = 0.5108$	0.5108	$\sigma^2 + r \sigma_{\alpha\beta}^2 + ra \sigma_\beta^2$
$\alpha\beta$	$(a-1)(b-1) = 1$	$(AB)_{yy} = 0.0029$	0.0029	$\sigma^2 + r \sigma_{\alpha\beta}^2$
Experimental Error	$(r-1)(ab-1) = 27$	$E_{yy} = 0.6096$	0.02258	σ^2
Total	$(rab-1) = 39$			

expected mean squares are also given in Table 13. The estimate of experimental error is 0.02258 as determined by the ANOV.

F-tests performed on the data from the analysis of variance indicated that operator effect was significant, $\sigma_o^2 \neq 0$, and also the material effect was significant, $\sigma_a^2 \neq 0$. Replicate effect and the operator-material interaction were determined to be not significant.

Materials might be expected to be significant. Different fineness modulus (different gradations) were deliberately chosen. Since the material used was a natural sand of a heterogeneous nature, it is not unusual to find different size fractions having different absorptions.

It is interesting to note that operators were significant, indicating the sensitivity of the test method to this variable. The procedure involved requires judgment as to when a saturated surface dry condition has been obtained and it also involves some amount of skill on the part of the operator in placing 500.0 grams of material into the flask quickly. It should be noted at this point that only two operators were used in this study and therefore it is difficult to make generalities about the whole population of operators with confidence. This study does give an indication as to what might be expected.

With estimate of the variances components available, limits were determined. As before, both confidence limits and control limits were determined. To obtain an estimate of the precision of an operator the control limits were placed on the individual observations. The procedures followed in determining the control limits were the same as those used in determining the control limits for absorption data of Test Method C 127-59. For this case a pooled estimate of σ^2 was determined as 0.0190. Since $\sigma_o^2 = 0$ and $\sigma_{ao}^2 = 0$ these were used with σ^2 to determine a pooled estimate of σ^2 . See Appendix A for the calculations of the components and the tests of hypothesis.



The pooled estimate of the variance σ^2 equals 0.0190 with 37 degrees of freedom. The control limits were determined as noted below.

For $\alpha = 0.05$

$$\text{Control limits} = \pm 2.262 \sqrt{\frac{0.0190}{1}} = \pm 0.312\%$$

For $\alpha = 0.01$

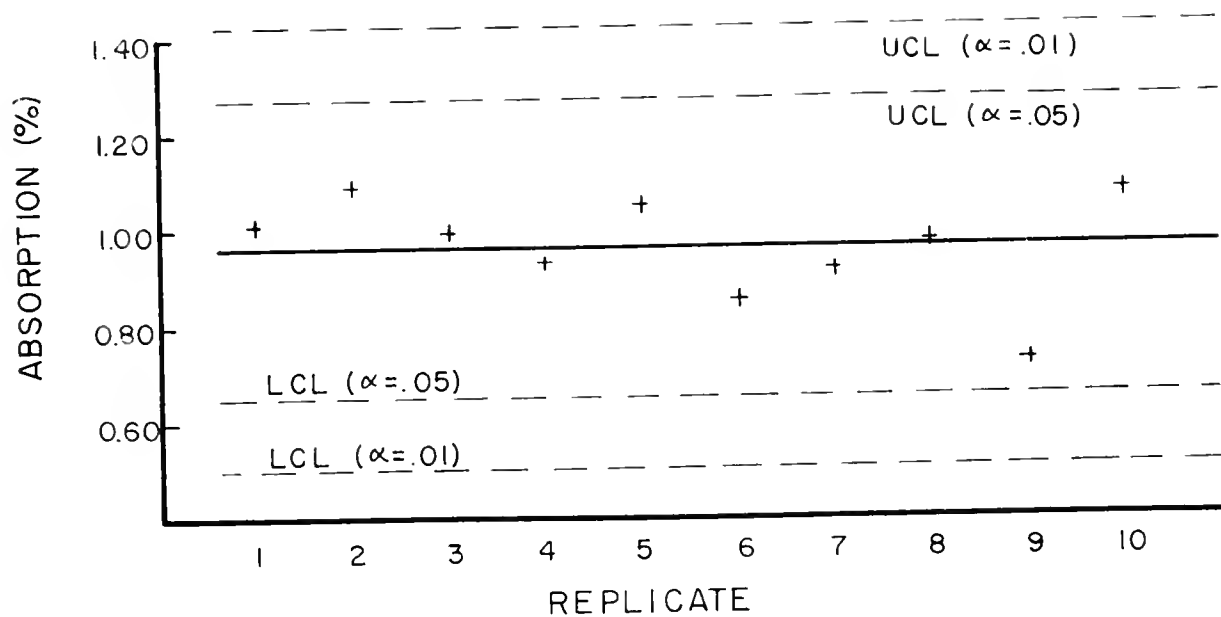
$$\text{Control limits} = \pm 3.25 \sqrt{\frac{0.0190}{1}} = \pm 0.448\%$$

Figures 9 and 10 shows these limits placed about the means of ten replicates.

Confidence limits were determined as follows. The pooled estimates of variance equal 0.0190, therefore σ equals 0.138. For an 0.05 α -level and 37 degrees of freedom, t equals 2.028. The 95% confidence limits are therefore, $\bar{Y} \pm 2.028 \frac{0.138}{\sqrt{n}} = \bar{Y} \pm \frac{0.280}{\sqrt{n}}$. For an 0.01 α -level and 37 degrees of freedom, t equals 2.716. The 99% confidence limits are therefore, $\bar{Y} \pm 2.716 \frac{0.138}{\sqrt{n}} = \bar{Y} \pm \frac{0.374}{\sqrt{n}}$. The interpretation of this is that we are 95% confident that the true mean, μ , lies with the limits estimated by $\bar{Y} \pm \frac{0.280}{\sqrt{n}}$ and that we are 99% confident that the true mean, μ , lies within the limits estimated by $\bar{Y} \pm \frac{0.374}{\sqrt{n}}$. \bar{Y} is the observed mean based on n observations.

Control limits were also determined for "between operators". These limits were determined calculating the variance and the degrees of freedom. The Satterwaite equation was used to determine the degrees of freedom, which for this case was one (see Appendix A for calculations). The control limits were determined as $\pm t \sqrt{\sigma^2_{\bar{Y}_j}}$. The $\sigma^2_{\bar{Y}_j}$ for this case was equal to

OPERATOR A



OPERATOR B

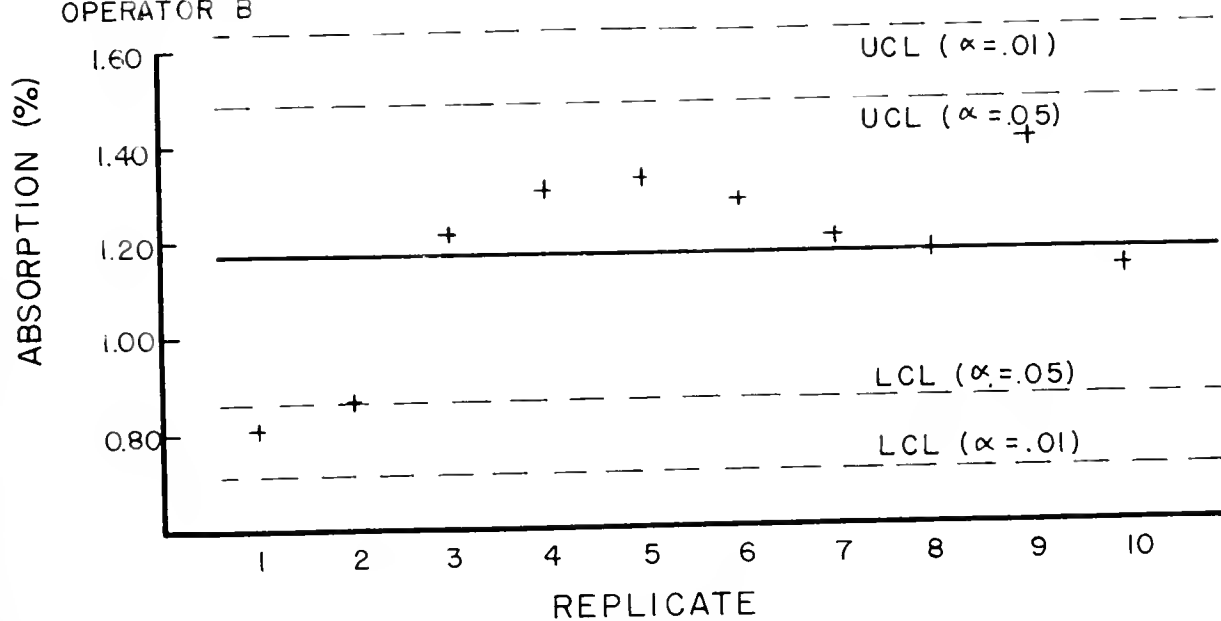


FIG. 9 CONTROL CHARTS FOR ABSORPTION DATA OF FINE AGGREGATE (F.M. = 2.40), TEST METHOD C 128-57.

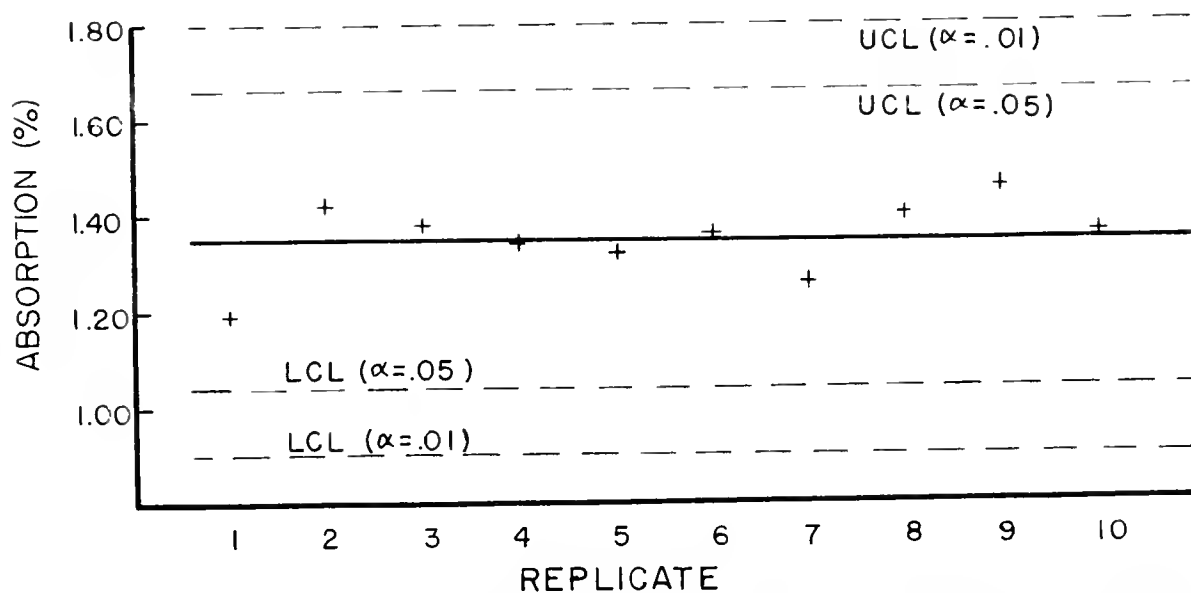
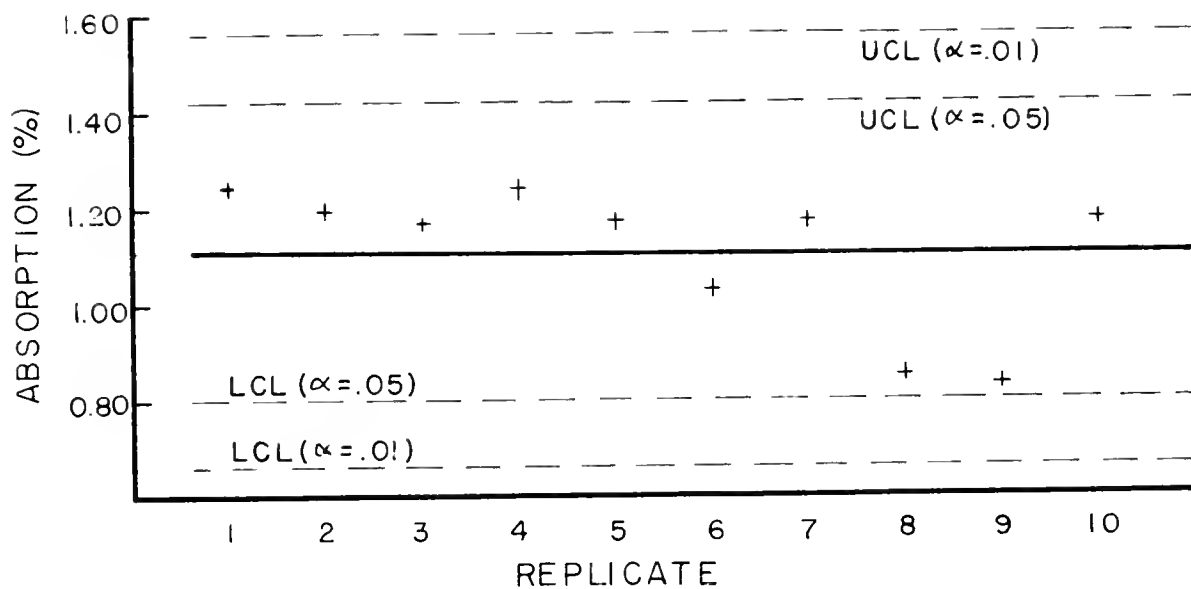


FIG. 10 CONTROL CHARTS FOR ABSORPTION DATA OF FINE AGGREGATE (F.M.= 3.00), TEST METHOD C 128-57.

0.01325. The limits were then $\pm t\sqrt{0.01325}$ where t was based on the chosen α -level and the degrees of freedom determined, which was one. For an α -level of 0.05 the control limits are $\pm 12.706 \sqrt{0.01325} = \pm 1.46\%$. These limits from a practical viewpoint are far too high. This points up, very dramatically the problems associated with trying to make generalities over a whole population with only one degree of freedom associated with the variance component due to operators. It should be further noted that the operator effect was determined significant, as previously mentioned, and this fact would tend to indicate that the control limits would be wider for this case.

A comparison of the error variances of this test method with that determined for Test Method C 127-59 shows that this test method has a much higher component of variance due to experimental error than did Test Method C 127-59. These error variances were 0.0226 and 0.00146, respectively.

Specific Gravity

Results of the analyses of variance performed on the data for the three types of specific gravity indicated that all have approximately the same magnitude of experimental error, namely $\sigma^2 = 0.0001$. However, the estimate of this was difficult since the magnitude of computational figures are small and the factor of significant figures reduces the precision of the analysis. This estimate of σ^2 however, if used in setting control limits on means of duplicate determination, replicates, gives ± 0.01 . This is the same as the practical limits determined for the specific gravities of coarse aggregate determined by Test Method C 127-59. The specific gravities of fine aggregate as determined by Test Method C 128-57 are more variable than those of coarse aggregates as determined by Test Method C 127-59. This fact may be noted by comparing the data of Table 8 with that of Table 12.

The variances associated with the specific gravity data were small but some very general conclusions may be made on the data. Namely, bulk specific gravity had more variance associated with it and is more sensitive to the operator variable and the material variable than does bulk specific gravity (SSD). Bulk gravity (SSD) in turn is more sensitive to the operator variable and the material variable than is apparent specific gravity. In fact, this study indicated no variance component due to the operator effect. Table 14 shows the analysis of variance for bulk specific gravity, Table 15 shows the ANOV for bulk specific gravity (SSD), and Table 16 shows the ANOV for apparent specific gravity.

General Observations

In the course of this study, it was noted that certain factors in Test Method C 128-57 could be improved or elaborated upon. The temperature bath specification contains no provision for any variance about 20°C. The time specified for the sample to remain in the bath also has no tolerance associated with it. The flask specified, 500 ml, is actually too small for the size of sample being tested.

The flask is not large enough to allow the sand to be mixed by the rolling action to remove entrapped air. It was noted that it was very difficult to remove the air entrapped in the sand at the bottom of the flask due to the fact that no mixture was achieved by the rolling action. A large flask would provide sufficient room for the sample to be mixed with the water and the entrapped air released.

No provision is made for the counter-action of foaming that may occur during the removal of air by the procedure specified.

TABLE 14
ANALYSIS OF VARIANCE TABLE FOR BULK SPECIFIC GRAVITY DATA

Source of Variation	Degrees of Freedom	Sum of Squares	Mean Square	Expected Mean Square
Replicates	$(r-1) = 9$	$R_{yy} = 0.0010$	0.000111	$\sigma^2 + ab\sigma_p^2$
α	$(a-1) = 1$	$A_{yy} = 0.0026$	0.0026	$\sigma^2 + r\sigma_{\alpha\beta}^2 + rb\sigma_{\alpha}^2$
β	$(b-1) = 1$	$B_{yy} = 0.0012$	0.0012	$\sigma^2 + r\sigma_{\alpha\beta}^2 + ra\sigma_{\beta}^2$
$\alpha\beta$	$(a-1)(b-1) = 1$	$(AB)_{yy} = 0$	0	$\sigma^2 + r\sigma_{\alpha\beta}^2$
Experimental Error	$(r-1)(ab-1) = 27$	$E_{yy} = 0.0030$	0.000111	σ^2
Total	$(rab-1) = 39$			

TABLE 15

ANALYSIS OF VARIANCE TABLE FOR BULK (SSD) SPECIFIC GRAVITY DATA

Source of Variation	Degrees of Freedom	Sum of Squares	Mean Square	Expected Mean Square
Replicates	$(r-1) = 9$	$R_{yy} = 0.0008$	0.00009	$\sigma^2 + ab\sigma_p^2$
α	$(a-1) = 1$	$A_{yy} = 0.0006$	0.00060	$\sigma^2 + r\sigma_{\alpha\beta}^2 + rb\sigma_\alpha^2$
β	$(b-1) = 1$	$B_{yy} = 0.0009$	0.00090	$\sigma^2 + r\sigma_{\alpha\beta}^2 + ra\sigma_\beta^2$
$\alpha\beta$	$(a-1)(b-1) = 1$	$(AB)_{yy} = 0$	0	$\sigma^2 + r\sigma_{\alpha\beta}^2$
Experimental Error	$(r-1)(ab-1) = 27$	$E_{yy} = 0.0015$	0.0000556	σ^2
Total	$(rab-1) = 39$			

TABLE 16
ANALYSIS OF VARIANCE TABLE FOR APPARENT SPECIFIC GRAVITY DATA

Source of Variation	Degrees of Freedom	Sum of Squares	Mean Square	Expected Mean Square
Replicates	$(r-1) = 9$	$R_{yy} = 0.0010$	0.00011	$\sigma^2 + ab \sigma_\rho^2$
α	$(a-1) = 1$	$A_{yy} = 0.0001$	0.00010	$\sigma^2 + r \sigma_{\alpha\beta}^2 + rb \sigma_\alpha^2$
β	$(b-1) = 1$	$B_{yy} = 0$	0	$\sigma^2 + r \sigma_{\alpha\beta}^2 + ra \sigma_\beta^2$
$\alpha\beta$	$(a-1)(b-1) = 1$	$(AB)_{yy} = 0$	0	$\sigma^2 + r \sigma_{\alpha\beta}^2$
Experimental Error	$(r-1)(ab-1) = 27$	$E_{yy} = 0.0019$	0.00007	σ^2
Total	$(rab-1) = 39$			

Table 17 gives the sieve analysis of the two samples after testing was completed.

Results Applicable to the Development
of Precision Statements

Both confidence limits and control limits were computed in the statistical analysis of the test data. Control limits define a range in values within which, in the long run, the sample statistics will fall a certain percentage of the time. Confidence limits define a range around the sample statistic that, with a given probability, will include the true value of the population parameter. Further, control limits indicate how well an operator is performing the test. If the operator's results are "out of control" some assignable cause may be the reason (e.g. the operator may not be following exactly the test procedures specified). For inclusion in a precision statement, control limits are superior to confidence limits since they provide more information relative to the actual performance of the test method.

There are many forms which a precision statement may take. The following are examples of several types that seem worthy of consideration for use.

1. A simple statement of the repeatability and reproducibility determined for the test method.

Example: Duplicate determinations should check to the nearest 0.02 in the case of a single operator and to the nearest 0.05 for between-operators.

2. A statement giving the estimate of experimental error determined for the test method, the estimate of each variance component, and the degrees of freedom associated with each.

TABLE 17
GRADATION OF FINE AGGREGATE AFTER TESTING

Sieve Fraction	Sample #1 (FM 2.40)	Sample #2 (FM 3.00)
3/8 in. - #4	0	0
#4 - #8	51.4 gm	96.1 gm
#8 - #16	222.6 gm	196.3 gm
#16 - #30	204.3 gm	224.7 gm
#30 - #50	197.9 gm	170.7 gm
#50 - #100	164.7 gm	179.3 gm
#100 - #200	22.9 gm	25.0 gm
Pan	<u>2.4</u> gm	<u>2.4</u> gm
	866.2 gm	894.5 gm

Example: The experimental error for the test method is 0.0015 with 27 degrees of freedom. The components of variance for replicates, operators, and materials are 0.0014 with 9 degrees of freedom, 0.008 with 1 degree of freedom, and 1.47 (highly significant) with 1 degree of freedom, respectively.

3. A statement giving the estimate of experimental error and the expression for computing control limits (e.g. $UCL = \bar{Y} + t \sqrt{\sigma_{\bar{Y}}^2}$, $LCL = \bar{Y} - t \sqrt{\sigma_{\bar{Y}}^2}$ where \bar{Y} = the sample mean and $\sigma_{\bar{Y}}^2$ = the variance about the mean). The degrees of freedom and the α -levels to be used with t must also be noted.

Example: The experimental error for the test method is 0.0015. Control limits are to be computed by $\bar{Y} \pm t \sqrt{\sigma_{\bar{Y}}^2}$, where $\sigma_{\bar{Y}}^2 = \sigma_{\epsilon}^2$ and t is based on an α -level of 0.05 and 9 degrees of freedom.

4. A statement giving the control limits for a single operator performing the test a number of times and the control limits for between-operators. The degrees of freedom and the α -level for determining t should also be given.

Example: Control limits for a single operator performing the test are $\bar{Y} \pm 0.03$, with 27 degrees of freedom associated with the estimate of experimental error and t based on 9 degrees of freedom and an α -level of 0.05. Control limits for between-operators are the same as for a single operator.

Example 4 provides the optimum information and seems to be the best for inclusion in test methods of the type considered in this study.

The results of this study indicate that it is essential to determine the variables for a test method and to develop a statistical model before actual testing begins. The estimate of error variance and other variance components can be determined from an analysis of variance. With this information available, significance test can then be made. Control limits

can be placed on the data of a single operator and on data for between-operators.

The results of the work on ASTM Designation: C 117-49 indicated that two sets of control limits for operators were required. Control limits computed for Operator A would be applicable to inexperienced personnel while the control limits computed for Operator B would be applicable to experienced laboratory personnel. The 95% and 99% control limits determined for the data of Operator A are $\pm 0.26\%$ and $\pm 0.38\%$, respectively, based on an error variance of 0.0134 and 9 degrees of freedom. The 95% and 99% control limits for the data of Operator B are $\pm 0.12\%$ and $\pm 0.16\%$, respectively, based on an error variance of 0.0031 and 19 degrees of freedom.

The results of the work on ASTM Designation: C 127-59 indicated that for the absorption data, operator-to-operator differences were insignificant and a single set of control limits for individual observations were determined. The 95% and 99% control limits are $\pm 0.09\%$ and $\pm 0.13\%$, respectively, based on an error variance of 0.00145 with 36 degrees of freedom and t based on the chosen α -level and 9 degrees of freedom.

In the case of specific gravity data, the variability was so small that an analysis of variance was unwarranted. Examination of the results led to the obvious conclusion that "practical limits" of ± 0.01 could be set for both the single operator and between-operator cases.

The absorption data collected using ASTM Designation: C 128-57 indicated that operator-to-operator variances were homogeneous yet means between operators were significantly different. This situation permits the calculation of a single set of control limits applicable to the

individual observations of any operator but also points out that the absorption value determined may differ between operators even when the data from both operators are "in control". The 95% and 99% control limits for individual observations are $\pm 0.31\%$ and $\pm 0.45\%$, respectively, based on an error variance of 0.0190 with 37 degrees of freedom, and t based on the chosen α -level and 9 degrees of freedom.

The control limits determined for "between-operators" were found to be far too high from a practical viewpoint. A study involving more operators is necessary to obtain a good estimate of the component of variance associated with operators. This test method is evidently quite sensitive to operator techniques.

In the case of specific gravity, the variability was very small, 0.0001. As in the case of C 127-59, examination of the results led to be obvious conclusion that practical limits of ± 0.01 could be set for both the "single" operator and the "between-operator" cases.

SUMMARY OF THE RESULTS AND CONCLUSIONS

The establishment of control limits requires the knowledge of each component of variance present in the design model. Three ASTM Test Methods were chosen for study. Sampling variation were eliminated by repeated testing of the same sample but other variables were incorporated in the model. Once the variables were determined a statistical model was chosen. Factorial design models were selected as being the best for this study.

The results of this study indicate that:

- a) the experimental designs used produce information useful in determining control limits applicable to individual observations of the attribute being measured.
- b) control limits on individual observations form the basis for useful precision statements.
- c) the data collected are also useful in the evaluation of the test method itself.

The control limits for individual observations determined for the data of C 117 and for the absorption data of C 127 and C 128 are wider than one would expect. C 117 contains no precision statement, but C 127 and C 128 state that duplicate determinations of absorption should check to the nearest 0.05%. The data collected were too widely scattered to permit the establishment of control limits as narrow as 0.05% even though the test method was followed in detail. Evidently some significant variable is not being controlled and if better reproducibility is desired, changes

in the test method are necessary. On the other hand, the specific gravity data of C 127 and C 128 had practically zero variance which indicates that the test methods as now written are adequate for this determination.

The control limits determined for means averaged over replicates (control limits for between-operators) were inconclusive. As stated previously it was not possible to determine these limits for Test Method C 117-49. For Test Method C 127-59 these limits were $\pm 0.0175\%$ and \pm to 0.0237% for 0.05 and 0.01 α -level respectively. For the Data of Test Method C 128-57 operators were found to be significant and control limits for between-operators were determined. (These limits were extremely large. The reason for this was that only one degree of freedom was associated with the determination of the estimate of variance associated with operators.) Based on these facts it can be said that more operators should be used in estimating this variance component. It is difficult to generalize about the whole population of operators with only one degree of freedom available in the analysis.

The following sections summarize the results and conclusions for each Test Method investigated.

C 117-49: Test for Material Finer than
the No. 200 Sieve in Aggregate

It was difficult to set control limits for this test method because non-homogeneity of variance was encountered due to the difference in the "character" of the operators. The results of the analysis made on the data for Operator B may be considered typical for application to situations where experienced laboratory technicians perform the test method. The control limits for this case are $\pm 0.12\%$ for an α -level of 0.05 and $\pm 0.16\%$ for an α -level of 0.01. The control limits computed for the data for

Operator A may be considered typical for application to situations where persons who are not trained laboratory technicians are performing the test method. These control limits are $\pm 0.26\%$ for an α -level of 0.05 and $\pm 0.38\%$ for an α -level of 0.01.

The limits which were computed should be used with caution since they are based on the results of one operator. However, with a comparison of the control limits determined by the two operators, one experienced and one inexperienced, a "typical" limit associated with the test method can be estimated. At the 0.05 α -level, for example, the limits are $\pm 0.12\%$ and $\pm 0.26\%$ for Operators B and A respectively. A typical value of the control limits would be $\pm 0.19\%$. Further investigation of the test method, using more operators, is needed to get a better estimate of the variance components.

The limits for Operator A are about twice as wide as those for Operator B. It may be noted, however, in Figure 5 that the variability tends to decrease as the replications increase. This indicates that as the operator performed the test he became more proficient. Even though a "break-in" period was used to familiarize the operator with the test method the more experienced he became the less variability occurred. This points up that when performing an investigation of this nature the operators doing the testing should be thoroughly experienced in the laboratory testing procedures prescribed in the test method. Control limits established would then be applicable to experienced personnel when performing the test method following exactly the procedures outlined in the test method. Inexperience will result in more variability.

If, in the light of further study, a variable component not considered in this study is found significant a revision of the control limits would

be necessary. If no other significant variable were found, a reduction in the precision of weight determinations could be made.

C 127-59: Test for Specific Gravity and

Absorption of Coarse Aggregate

Control limits may be set on the available data for absorption of aggregate by use of the equation $\bar{Y} \pm \sqrt{\frac{\sigma_Y^2}{1}}$ placed on individual observations. The control limits are $\pm 0.09\%$ for an α -level of 0.05 and $\pm 0.13\%$ for an α -level of 0.01.

The variance of the specific gravity results, for each type of specific gravity (bulk, bulk saturated surface dry, and apparent), was found to be essentially zero. An examination of the data indicates that practical limits of ± 0.01 may be used. This applies to between-operators as well as to a single operator.

Between-operator control limits for the absorption data were found to be narrower than the control limits for a single operator. They are $\pm 0.0175\%$ at the 0.05 α -level and $\pm 0.0237\%$ at the 0.01 α -level.

C 128-57: Test for Specific Gravity and

Absorption of Fine Aggregate

The results of the analysis on the data for absorption indicated that operators and materials were significant. As a result of the significance of the two components the control limits were placed about means averaged over replicates. The control limits for individual observations are $\pm 0.31\%$ for an α -level of 0.05 and $\pm 0.45\%$ for an α -level of 0.01.

The absorption data show a high variability. This high variability may be the result of some assignable cause. One such cause could be

differences in laboratory humidity during the testing period. Another could be the result of the operator's failure to watch their samples closely during the drying period to determine when the saturated-surface dry condition was reached. If, in the light of further study, a variable not controlled in this investigation is found significant an adjustment in the limits would be necessary.

Control limits for between-operators were extremely large, $\pm 1.46\%$. The cause of this high value seems to be the fact that only one degree of freedom is associated with the estimate of the component of variance associated with operators. More operators should be used to obtain a better estimate of this component of variance.

Control limits computed for the specific gravity data were ± 0.01 for all three types of specific gravities. The between-operators control limits were found to be the same, ± 0.01 . These control limits are the same as the "practical" limits determined for Test Method C 127-59 and compare to the repeatability value of 0.02 given in both Test Methods C 127-59 and C 128-57.

Precision Statements for Test Methods Studied

On the basis of this study the following precision statements are suggested. These statements are based on the previous assumptions and include components of variance due to replicates, operators, and material. This study specifically eliminated the component of variance due to sampling and other variables not mentioned. The precision statements are for test data when the tests are performed by experienced personnel following exactly the procedures as outlined in the ASTM Standards for the test method.

Test Method C 117-49: Test for Material Finer

than the No. 200 Sieve in Aggregate

Control limits for a single operator performing the test are $\bar{Y} \pm 0.12$ percent, with 63 degrees of freedom associated with the estimate of experimental error and t based on 9 degrees of freedom and an α -level of 0.05.

Test Method C 127-59: Test for Specific Gravity

and Absorption of Coarse Aggregate

Absorption. Control limits for a single operator performing the test are $\bar{Y} \pm 0.09$ percent, with 27 degrees of freedom associated with the estimate of experimental error and t based on 9 degrees of freedom and an α -level of 0.05. Control limits for between-operators are the same.

Specific Gravity. Practical control limits for both a single operator and between-operators are $\bar{Y} \pm 0.01$, based on 40 observations.

Test Method C 128-57: Test for Specific Gravity

and Absorption of Fine Aggregate

Absorption. Control limits for a single operator performing the test are $\bar{Y} \pm 0.31$ percent, with 27 degrees of freedom associated with the estimate of experimental error and t based on 9 degrees of freedom and an α -level of 0.05.

Specific Gravity. Practical control limits for both a single operator and between-operators are $\bar{Y} \pm 0.01$, based on 40 observations.

PROPOSED RESEARCH

In light of information obtained in this study it is recommended that further research in this area take the following form:

1. Classification of Test Methods as to need for precision statement.
2. Investigation of each test method to determine what variables are present and what type of statistical model is best suited for the analysis of the method.
3. Analysis of test methods including sampling as a variable and other variables such as laboratory temperature and humidity.
4. Formulation of precision statements for all test methods which will benefit from them.

BIBLIOGRAPHY

BIBLIOGRAPHY

1. ASTM Committee E-11 on Quality Control of Materials, Recommended Procedure for Conducting on Interlaboratory Study of a Test Method, June, 1961.
2. ASTM Manual on Quality Control of Materials, Special Technical Publication 15-C, American Society for Testing Materials, 1951.
3. ASTM Standards 1958, Part 4.
4. ASTM Proceedings, p. 379, Appendix IV, 1953.
5. Brown, A. B., Precision of Present ASTM Tests on Bitumens and Bituminous Materials, ASTM Special Technical Publication No. 252.
6. Burr, I. W., Engineering Statistics and Quality Control, McGraw-Hill Book Company, Inc., New York, 1953, Chapter 15.
7. Cochran, W. G., Some Consequences when the Assumptions for Analysis of Variance are Not Satisfied, Biometrics, Vol. 3, No. 1, March, 1947.
8. Conner, W. S., How to Evaluate Precision, Material Research and Standards, Vol. 1, No. 4, April, 1961, pp. 272.
9. Crandall, J. R. and Blaine, R. L., Statistical Evaluation of Interlaboratory Cement Tests, ASTM Proceedings, Vol. 59, 1959, pp. 1129.
10. Eisenhart, Churchill, The Assumptions Underlying the Analysis of Variance, Biometrics, Vol. 3, No. 1, March 1947.
11. Kaplan, M. F., The Reproducibility of Flexure and Compression Tests for Determining the Strength of Cement, ASTM Proceedings, Vol. 60, 1960, pp. 1006.
12. Ostle, B., Statistics in Research, The Iowa State College Press, Ames, Iowa, 1954.
13. Report of Committee D-6 on Paper Products, Sixty-third Annual Meeting of the American Society for Testing Materials, 1960.
14. Report to Subcommittee II-a of ASTM Committee C-9 on the Present Status and Need for Precision Statements in Methods of Test, 1960.
15. Steel, R. G. D. and Torrie, J. H., Principles and Procedures of Statistics, McGraw-Hill Book Company, Inc., New York, 1960.

16. Supplement to Book of ASTM Standards, Part 4, 1959.
17. Terry, M. E., On the Analysis of Planned Experiments, Materials Research and Standards, Vol. 1, No. 4, April, 1961, pp. 273.
18. Youden, W. J., Experimental Designs and ASTM Committees, Materials Research and Standards, Vol. 1, No. 11, November, 1961, p. 862.
19. Youden, W. J., How to Evaluate Accuracy, Materials Research and Standards, Vol. 1, No. 4, April, 1961, p. 272.
20. Youden, W. J., Statistical Aspects of the Cement Testing Program, ASTM Proceedings, Vol. 59, 1959, p. 1120.

APPENDIX

Statistical Analysis of Test Data

Bartlett Test For Homogeneity of Variance
(Operator A for C 117-49)

	Limestone				Clay			
	1%	4%	7%	10%	1%	4%	7%	10%
X.	10.66	40.52	70.06	100.30	10.80	39.12	69.56	99.88
\bar{X}	1.066	4.052	7.006	10.030	1.080	3.912	6.956	9.988
ΣX_i^2	11.4052	164.3112	490.9708	1006.0924	11.7576	153.3824	483.9464	997.6632
$\frac{\Sigma X_i^2}{r}$	11.3636	164.1870	490.8404	1006.0090	11.6640	153.0374	483.8594	997.6014
SS	0.0416	0.1242	0.1304	0.0834	0.0936	0.3450	0.0870	0.0618
df	9	9	9	9	9	9	9	9
S_j^2	0.004622	0.013800	0.014489	0.009267	0.010400	0.03833	0.009667	0.006867
$\log S_j^2$	-2.33517	-1.86012	-1.83896	-2.03306	-1.98297	-1.41642	-2.01471	-2.16323
$(r-1) \log S_j^2$	-21.01828	-16.74108	-16.55064	-18.29754	-17.84673	-12.74778	-18.13239	-19.46904

$$\Sigma (r-1) \log S_j^2 = -140.80347$$

$$\Sigma SS = 0.9670 \quad S_j^2 = \frac{0.9670}{72} = 0.01343$$

$$\log S_j^2 = -1.87192 \quad ab(r-1) \log S_j^2 = -1.87192 \times 72 = -134.77823$$

$$2.3026 [-134.77824 - (-140.80347)] = 2.3026 \times 6.02523 = 14.48$$

$$C = 1 + \frac{1}{3(K-1)} \left[\frac{\Sigma \frac{1}{(r-1)}}{\Sigma \frac{1}{(r-1)}} \right]$$

$$C = 1 + \frac{1}{3 \times 7} \left[\frac{8}{9} - \frac{1}{72} \right] = 1 + \frac{1}{21} \left[\frac{64}{72} - \frac{1}{72} \right] = 1 + \frac{63}{21 \times 72}$$

$$C = 1 + \frac{1}{24} = 1 + 0.04167 = 1.04167$$

$$\chi^2_{\text{corr.}} = \frac{14.48}{1.04167} = 13.90 \text{ (corrected } \chi^2 \text{ with 7df)}$$

$\chi^2_{.05, 7\text{df}} = 14.1 > 13.90$, therefore accept H_0 : all variances equal at the .05 α -level.

Bartlett Test for Homogeneity of Variance
(Operator B for C 117-49)

Limestone						Clay			
	1%	4%	7%	10%	1%	4%	7%	10%	
X.	10.24	39.74	69.94	99.84	10.04	39.96	69.90	100.00	
\bar{X}	1.024	3.974	6.994	9.984	1.004	3.996	6.990	10.000	
ΣX_i^2	10.4952	157.9468	489.2196	996.8432	10.1040	159.7032	488.6276	1000.0200	
$\frac{\Sigma X_i^2}{n}$	10.4858	157.9268	489.1604	996.8026	10.0802	159.6802	488.6010	1000.0000	
SS	0.0094	0.0200	0.0592	0.0406	0.0238	0.0230	0.0266	0.0200	
df	9	9	9	9	9	9	9	9	
S_j^2	0.001044	0.002222	0.006578	0.004511	0.002644	0.002556	0.002956	0.002222	
$\log S_j^2$	-2.98130	-2.65326	-2.18191	-2.34573	-2.57774	-2.59244	-2.52930	-2.65326	
$(r-1) \log S_j^2$	-26.83170	-23.87934	-19.63719	-21.11157	-23.19966	-23.33196	-22.76370	-23.87934	

$$\Sigma (r-1) \log S_j^2 = -184.63446$$

$$\Sigma SS = 0.2226 \quad S_j^2 = \frac{0.2226}{72} = 0.0030917$$

$$\log S_j^2 = -2.50980 \quad ab(r-1) \log S_j^2 = -2.50980 \times 72 = -180.70560$$

$$2.3026 [-180.70560 - (-184.63446)] = 2.3026 \times 3.92886 = 9.0466$$

$$C = 1 + \frac{1}{3(K-1)} \left[\Sigma \frac{1}{n_i-1} - \frac{1}{\Sigma(n-1)} \right]$$

$$C = 1 + \frac{1}{3 \times 7} \left[\frac{8}{9} - \frac{1}{72} \right] = 1 + \frac{1}{21} \left(\frac{64}{72} - \frac{1}{72} \right) = 1 + \frac{63}{21 \times 72}$$

$$C = 1 + \frac{1}{24} = 1 + 0.04167 = 1.04167$$

$$\frac{9.0466}{1.04167} = 8.70, \text{ corrected } \chi^2 \text{ with 7df}$$

$\chi^2_{.05, 7df} = 14.1 > 8.70$ therefore, we may accept H_0 : that all variances are equal at the .05 α -level.

Bartlett Test for Homogeneity of Variance
(Operators A and B for C 117-49)

$$\Sigma (r-1) \log S_j^2 = -140.80347 - 184.63446 = -325.43793$$

$$\Sigma SS = 0.9670 + 0.2226 = 1.1896$$

$$S_j^2 = \frac{1.1896}{144} = 0.00826$$

$$\log S_j^2 = -2.08297$$

$$ab (r-1) \log S_j^2 = 144 (-2.08297) = -299.94768$$

$$2.3026 [-299.94768 - (-325.43793)] = 58.7$$

$$C = 1 + \frac{1}{3(K-1)} \left[\Sigma \frac{1}{r-1} - \frac{1}{\Sigma(r-1)} \right]$$

$$C = 1 + \frac{1}{45} \left[\frac{16}{9} - \frac{1}{144} \right] = 1.0393$$

$$\chi^2_{\text{corr.}} = \frac{58.7}{1.0393} = 56.5$$

$$\chi^2_{.05, 15\text{df}} = 25.0 < 56.5 \quad \therefore \text{significant}$$

$$\chi^2_{.01, 15\text{df}} = 30.6 < 56.5 \quad \therefore \text{significant}$$

Therefore reject H_0 : that all variances are equal, at the .05 and .01 α -level.

Calculations of Sum of Squares
(Operator A for C 117-49)

TOTALS OF 10 REPLICATES (axb Table)

Levels	<u>Material</u>		Totals
	LS	CL	
1%	10.66	10.80	21.46
4%	40.52	39.12	79.64
7%	70.06	69.52	139.62
10%	100.30	99.88	200.18
Totals	221.54	219.36	440.90

α = Mat.

β = levels

r = 10 (reps)

a = 2 (mat.)

b = 4 (amount)

$$\begin{aligned} \Sigma_y^2 &= \text{total sum of squares} = \sum_{i=1}^{i=r} \sum_{j=1}^{j=2} \sum_{k=1}^{k=4} y_{ijk}^2 - \frac{T^2}{rab} \\ &= 3319.5292 - \frac{(440.90)^2}{80} = 889.6191 \end{aligned}$$

$$\begin{aligned} R_{yy} &= \text{Replicate sum of squares} = \frac{\sum_{i=1}^{i=n} R_i^2}{ab} - \frac{T^2}{rab} \\ &= 2430.0920 - \frac{(440.90)^2}{80} = 0.1819 \end{aligned}$$

$$\begin{aligned} T_{yy} &= \text{treatment sum of squares} = \frac{\sum_{j=1}^{j=2} \sum_{k=1}^{k=4} T_{jk}^2}{r} - \frac{T^2}{rab} \\ &= 3318.5622 - \frac{(440.90)^2}{80} = 888.6521 \end{aligned}$$

A_{yy} = Sum of squares associated with the different levels of a

$$A_{yy} = \frac{\sum_{j=1}^{j=a} A_j^2}{rb} - \frac{T^2}{rab} = \frac{(221.54)^2 + (219.36)^2}{40} - \frac{(440.90)^2}{80} = 0.0594$$

B_{yy} = Sum of squares associated with the different levels of b

$$\begin{aligned} B_{yy} &= \frac{\sum_{k=1}^{k=4} B_k^2}{ra} - \frac{T^2}{rab} = \frac{(21.46)^2 + (79.64)^2 + (139.62)^2 + (200.18)^2}{20} - \frac{(440.90)^2}{80} \\ &= 888.5318 \end{aligned}$$

$(AB)_{yy}$ = Sum of squares associated with the interaction of the two factors a and b.

$$= T_{yy} - A_{yy} - B_{yy} = 0.0609$$

E_{yy} = experimental error sum of squares

$$= \sum_y^2 - R_{yy} - T_{yy} = 0.7851$$

where:

T = total of all obs.

R_i = total of all obs. in the i th rep.

T_{jk} = entry in the (jk) th cell of the axb table, this entry being the total of all obs. associated with the j th level of factor a and the k th level of factor b.

A_j = total of all obs. associated with the j th level of factor a.

B_k = total of all obs. associated with the k th level of factor b.

Calculation of the Components of Variance for C 117-49

Operator A

$$\sigma^2 = 0.0124$$

$$\sigma^2 + r \sigma_{\alpha\beta}^2 = 0.0203$$

$$\sigma_{\alpha\beta}^2 = \frac{0.0203 - 0.0124}{10} = 0.0008$$

$$\sigma^2 + rb \sigma_{\alpha}^2 = 0.0594$$

$$rb \sigma_{\alpha}^2 = 0.0594 - 0.0124$$

$$\sigma_{\alpha}^2 = \frac{0.0470}{40} = 0.0012$$

$$\sigma^2 + ab \sigma_{\rho}^2 = 0.0202$$

$$ab \sigma_{\rho}^2 = 0.0202 - 0.0124$$

$$\sigma_{\rho}^2 = \frac{0.0078}{8} = 0.0010$$

Operator B

$$\sigma^2 = 0.0028$$

$$\sigma_{\alpha\beta}^2 = 0$$

$$\sigma_{\alpha}^2 = 0$$

$$\sigma_{\rho}^2 = \frac{0.0054 - 0.0028}{8} = 0.0003$$

Tests of Hypotheses (Operator A for C 117-49)

$$H_1: \sigma_{\alpha\beta}^2 = 0 \text{ (Amount - Operator Interaction Effect)}$$

$$\frac{MS_{\alpha\beta}}{MS_E} = \frac{0.0203}{0.0124} = 1.64 < F_{.05, 3 \text{ \& } 63df} = 2.75$$

Therefore accept H_1 .

$$H_2: \beta_k = 0 \text{ (Operator effect)}$$

$$\frac{MS_{\beta}}{MS_{\alpha\beta}} = \frac{296.1773}{0.0203} = 14,600 > F_{.05, 3, 63df} = 2.75$$

Therefore we are unable to accept H_2 .

$$H_3: \sigma_{\alpha}^2 = 0 \text{ (Amount effect)}$$

$$\frac{MS_{\alpha}}{MS_E} = \frac{0.0594}{0.0124} = 4.79 > F_{.05, 1, 63df} = 3.99$$

Therefore we are unable to accept H_3 .

$$H_4: \rho_1 = 0 \text{ (Replicate effect)}$$

$$\frac{MS_{\rho}}{MS_E} = \frac{0.0202}{0.0124} = 1.63 < F_{.05, 9, 63df} = 2.02$$

Therefore accept H_4 .

Tests of Hypotheses (Operator B for C 117-49)

$$H_1: \sigma_{\alpha\beta}^2 = 0 \text{ (Amount-Operator Interaction Effect)}$$

$$\frac{MS_{\alpha\beta}}{MS_E} = \frac{0.0025}{0.0028} = 0.89 < F_{.05} \quad 3, 63df = 2.75$$

Therefore accept H_1 .

$$H_2: \beta_k = 0 \text{ (Operator effect)}$$

$$\frac{MS_{\beta}}{MS_{\alpha\beta}} = \frac{290.8232}{0.0025} = 119,500 > F_{.05} \quad 3, 63df = 2.75$$

Therefore we are unable to accept H_2 .

$$H_3: \sigma_a^2 = 0 \text{ (Amount effect)}$$

$$\frac{MS_a}{MS_E} = \frac{0.0003}{0.0028} = 0.11 < F_{.05} \quad 1, 63df = 3.99$$

Therefore accept H_3 .

$$H_4: \rho_i = 0 \text{ (Replicate effect)}$$

$$\frac{MS_{\rho}}{MS_E} = \frac{0.0054}{0.0028} = 1.93 < F_{.05} \quad 9, 63df = 2.02$$

Therefore accept H_4 .

Bartlett Test for Homogeneity of Variance
(Absorption for C 127-59)

	Limestone		Gravel	
	Operator A	Operator B	Operator A	Operator B
X.	8.88	8.95	13.08	12.43
\bar{X}	0.888	0.895	1.308	1.243
ΣX_i^2	7.8942	8.0199	17.1330	15.4599
$\frac{\Sigma X_i^2}{r}$	7.8854	8.0102	17.1086	15.4505
SS	0.0088	0.0097	0.0244	0.0094
df	9	9	9	9
S_j^2	0.00098	0.00108	0.00271	0.00104
$\log S_j^2$	-3.00877	-2.96658	-2.56703	-2.98297
$(r-1) \log S_j^2$	-27.07893	-26.69922	-23.10327	-26.84673

$$\Sigma (r-1) \log S_j^2 = -103.72815$$

$$\Sigma SS = .0523 \quad S_j^2 = \frac{.0523}{36} = .00145$$

$$\log S_j^2 = -2.83863 \times 36 = -102.19068$$

$$2.3026 [-102.19068 - (-103.72815)] = 2.3026 \times 1.53747 = 3.84$$

$$C = 1 + \frac{1}{3(k-1)} \left[\Sigma \frac{1}{n_i-1} - \frac{1}{\Sigma n-1} \right]$$

$$C = 1 + \frac{1}{3 \times 3} \left[\frac{4}{9} - \frac{1}{36} \right] = 1 + \frac{1}{9} \left(\frac{15}{36} \right) = 1 + \frac{15}{324}$$

$$C = 1.0463$$

$$\frac{3.84}{1.0463} = 3.67 = \text{Corrected } \chi^2 \text{ with } 3(df)$$

$\chi_{.05}^2$, 3df = 7.81 > 3.67 Therefore we may accept the hypothesis

that variances are equal at .05 α -level.

Calculations for Sum of Squares
(Absorption for C 127-59)

axb Table

Operator	Material		Totals
	LS	CL	
A	8.88	13.08	21.96
B	8.95	12.43	21.38
Totals	17.83	25.51	43.34

$$\frac{T^2}{rab} = \frac{(43.34)^2}{40} = 46.9589$$

$$\sum \sum Y_{ijk}^2 = 48.5070$$

$$\frac{\sum R_i^2}{ab} = \frac{187.8868}{4} = 46.9717$$

$$\frac{\sum T_{jk}^2}{r} = 48.4548$$

$$\frac{\sum A_j^2}{rb} = \frac{968.6690}{20} = 48.4334$$

$$\frac{\sum B_k^2}{ra} = \frac{939.3460}{20} = 46.9673$$

$$\Sigma_y^2 = 48.5070 - 46.9589 = 1.5481$$

$$R_{yy} = 46.9717 - 46.9589 = 0.0128$$

$$S_{ab} = 48.4548 - 46.9589 = 1.4959$$

$$A_{yy} = 48.4334 - 46.9589 = 1.4745$$

$$B_{yy} = 46.9673 - 46.9589 = 0.0084$$

$$(AB)_{yy} = S_{ab} - A_{yy} - B_{yy} = 0.0130$$

$$E_{yy} = \Sigma_y^2 - R_{yy} - S_{ab} = 0.0394$$

Calculation of the Components of Variance
(Absorption for C 127-59)

$$\sigma_{\alpha\beta}^2 = \frac{.01300 - 0.00146}{10} = \frac{.01254}{10} = 0.00125$$

$$\sigma_{\beta}^2 = \frac{.00840 - 0.01300}{20} = -0.0046 = 0$$

σ_{α}^2 of no real interest since materials are generally expected to have different absorption.

$$\sigma_r^2 = \frac{.00142 - 0.00146}{4} = 0$$

$$\sigma_E^2 = 0.00146$$

Tests of Hypotheses (Absorption for C 127-59)

$H_1: \sigma^2_{\alpha\beta} = 0$ (Material-Operator Interaction effect)

$$\frac{MS_{\alpha\beta}}{MS_E} = \frac{0.01300}{0.00146} = 8.90 > F_{.05} \quad 1, 27df = 4.21$$

Therefore we are unable to accept H_1 .

$H_2: \sigma^2_{\beta} = 0$ (Operator effect)

$$\frac{MS_{\beta}}{MS_{\alpha\beta}} = \frac{0.0084}{0.0130} = 0.65 < F_{.05}, \quad 1, 1df = 161$$

Therefore accept H_2 .

$H_3: \sigma^2_{\alpha} = 0$ (Material effect)

$$\frac{MS_{\alpha}}{MS_{\alpha\beta}} = \frac{1.4745}{0.0130} = 113.42 < F_{.05}, \quad 1, 1df = 161$$

Therefore accept H_3 .

$H_4: \sigma^2_e = 0$ (Replicate effect)

$$\frac{MS}{MS_E} = \frac{0.00142}{0.00146} = 0.97 < F_{.05}, \quad 9, 27df = 2.25$$

Therefore accept H_4 .

Calculation of Control Limits for Between-Operators
(Absorption for C 127-59)

$$V(\bar{Y}_j) = \frac{\sigma_e^2}{10} + \frac{\sigma_\theta^2}{2} + \frac{\sigma_E^2}{20}$$

$$V(\bar{Y}_j) = \frac{\sigma_e^2}{10} + \frac{\sigma_\theta^2}{2} + \frac{\sigma^2}{20} = 0 + 0 + \frac{0.00146}{20} = 0.00073$$

$$\pm t \sqrt{.000073} = \pm t \ 0.00854$$

$$\text{df for } t = 26 = (n-1) \text{ df for } \sigma_E^2$$

$$\alpha = 0.05$$

$$\pm t \ 0.0121 = 2.052 \times 0.00854 = \pm 0.0175$$

$$\alpha = 0.01$$

$$\pm t \ 0.0121 = 2.771 \times 0.00854 = \pm 0.0237$$

These limits are narrower than the limits on individual observations. Since the components of variance due to operators, σ_θ^2 , and replicates, σ_e^2 , were found to equal zero the degrees of freedom used was 27, that associated with σ_E^2 .

Bartlett Test for Homogeneity of Variance
(Absorption for C 128-57)

	FM 2.40		FM 3.00	
	Operator A	Operator B	Operator A	Operator B
X.	9.60	11.69	11.06	13.49
\bar{X}	0.960	1.169	1.106	1.349
$\sum X_i^2$	9.3250	13.9937	12.4392	18.2533
$\frac{\sum x_i^2}{r}$	9.2160	13.6656	12.2324	18.1980
SS	0.1090	0.3381	0.2068	0.1553
df	9	9	9	9
S_j^2	0.01211	0.03757	0.02298	0.01726
$\log S_j^2$	- 1.91686	- 1.42504	- 1.63865	- 1.76296
$(r-1) \log S_j^2$	-17.25174	-12.82536	-14.74785	-15.86664

$$\sum (r-1) \log S_j^2 = - 60.69189$$

$$\sum SS = .8092 \quad S_j^2 = \frac{.8092}{36} = 0.02248$$

$$\log S_j^2 = - 1.64820 \times 36 = - 59.33520$$

$$2.3026 [- 59.33520 - (- 60.69189)] = 2.3026 \times 1.35669 = 3.1239$$

$$C = 1.0463$$

$$\chi_{\text{corr.}}^2 = \frac{3.1239}{1.0463} = 2.9885, 3df$$

$$\chi_{.05}^2, 3df = 7.81$$

Since $2.99 < 7.81$ accept H_0 : that the variances are equal at the
0.05 α -level.

Calculations for Sum of Squares
(Absorption for C 128-57)

axb Table

Operator	Material		Totals
	2.40	3.00	
A	9.60	11.06	20.66
B	11.69	13.49	25.18
Totals	21.29	24.55	45.84

$$\frac{T^2}{rab} = \frac{(45.84)^2}{40} = \frac{2101.3056}{40} = 52.5326$$

$$\sum \sum Y_{ijk}^2 = 54.0112$$

$$\frac{\sum R_i^2}{ab} = \frac{210.4890}{4} = 52.6222$$

$$\frac{\sum T_{jk}^2}{r} = \frac{533.1198}{10} = 53.31198$$

$$\frac{\sum A_j^2}{rb} = \frac{1055.9666}{20} = 52.7983$$

$$\frac{\sum B_k^2}{ra} = \frac{1060.8680}{20} = 53.0434$$

$$\Sigma_y^2 = 54.0112 - 52.5326 = 1.4786$$

$$R_{yy} = 52.6222 - 52.5326 = 0.0896$$

$$S_{ab} = 53.3120 - 52.5326 = 0.7794$$

$$A_{yy} = 52.7983 - 52.5326 = 0.2657$$

$$B_{yy} = 53.0434 - 52.5326 = 0.5108$$

$$(AB)_{yy} = S_{ab} - A_{yy} - B_{yy} = 0.0029$$

$$E_{yy} = \Sigma_y^2 - R_{yy} - S_{ab} = 0.6096$$

Calculations of the Components of Variance
(Absorption for C 128-57)

$$\sigma_r^2 = 0$$

$$\sigma_{\alpha\beta}^2 = 0$$

$$\sigma_E^2 = 0.0226$$

$$\text{Pooled } S_E^2 = \frac{0.0896}{9} + \frac{0.0029}{1} + \frac{0.6096}{27} = \frac{0.7021}{37} = 0.0190$$

$$\sigma_\alpha^2 = \frac{0.2657 - \sigma_E^2 \text{ pooled}}{rb} = \frac{0.2657 - 0.0190}{20} = 0.0123$$

$$\sigma_\beta^2 = \frac{0.5108 - \sigma_E^2 \text{ pooled}}{ra} = \frac{0.5108 - 0.0190}{20} = 0.0246$$

In testing H_2 : $\sigma_r^2 = 0$, the MS_E and the $MS_{\alpha\beta}$ were pooled to form a better estimate of σ^2 . Since H_2 was accepted the MS_E , $MS_{\alpha\beta}$, and the MS were pooled to form a better estimate of σ^2 . This pooled MS was then used to test H_3 and H_4 .

Tests of Hypotheses (Absorption for C 128-57)

$H_1: \sigma_{\alpha\beta}^2 = 0$ (Material-Operator Interaction effect)

$$\frac{MS_{\alpha\beta}}{MS_E} = \frac{0.0029}{0.0226} = 0.13 < 4.21 = F_{.05, 1, 27df}$$

Therefore accept H_1 .

$H_3: \sigma_{\beta}^2 = 0$ (Operator effect)

$$\frac{MS_{\beta}}{MS_E \text{ pooled}} = \frac{0.5108}{0.0190} = 26.90 > 4.10 = F_{.05, 1, 37df}$$

Therefore we are unable to accept H_3 .

$H_4: \sigma_{\alpha}^2 = 0$ (Material effect)

$$\frac{MS_{\alpha}}{MS_E \text{ pooled}} = \frac{0.2657}{0.0190} = 13.98 > 4.10 = F_{.05, 1, 37df}$$

Therefore we are unable to accept H_4 .

$H_2: \sigma_{\epsilon}^2 = 0$ (Replicate effect)

$$\frac{MS}{MS_E \text{ pooled}} = \frac{0.00995}{0.0218} = 0.46 < 2.24 = F_{.05, 9, 28df}$$

Therefore accept H_2 .

Calculations of Control Limits for Between-Operators
(Absorption for C 128-57)

$$v(\bar{Y}_j) = \frac{\sigma_e^2}{10} + \frac{\sigma_e^2}{2} + \frac{\sigma_e^2}{20} = 0 + 0.0123 + \frac{0.0190}{20} = 0.01325$$

$$f' = \frac{[v_1 + v_5 + v_3 + v_4]^2}{\frac{v_1^2}{9} + \frac{v_5^2}{27} + \frac{v_3^2}{1} + \frac{v_4^2}{1}}$$

$$f' = \frac{[0.00995 + 0.02258 + 0.51080 + 0.00290]^2}{\frac{(0.00995)^2}{9} + \frac{(0.02258)^2}{27} + \frac{(0.51080)^2}{1} + \frac{(0.00290)^2}{1}} = 1$$

The control limits are $\bar{Y} \pm t_{.05, 1df} \sqrt{0.01325} = \bar{Y} \pm 1.46\%$.

These limits from a practical viewpoint are far too high. This points up very dramatically the problem associated with trying to make generalities over a population with only one degree of freedom associated with the variance component due to operators.



Bartlett Test for Homogeneity of Variance
(Bulk Specific Gravity for C 128-57)

	FM = 2.40		FM = 3.00	
	Operator A	Operator B	Operator A	Operator B
X.	26.15	25.98	20.03	25.88
\bar{X}	2.615	2.598	2.603	2.588
ΣX_i^2	68.3833	67.49680	67.75770	66.978000
$\frac{\Sigma X_i^2}{r}$	68.38225	67.49604	67.75609	66.97744
SS	0.00105	0.00076	0.00151	0.00056
df	9	9	9	9
S_j^2	0.00012	0.00008	0.00017	0.00006
$\log S_j^2$	- 3.92082	- 4.09691	- 3.76555	- 4.22185
$(r-1) \log S_j^2$	-35.28738	-36.87219	-33.88995	-37.99665

$$\Sigma (r-1) \log S_j^2 = - 144.04617$$

$$\Sigma SS = 0.00398 \quad \frac{0.00398}{36} = 0.00011$$

$$\log S_j^2 = - 3.95861 \times 36 = - 142.50996$$

$$1.53621 \times 2.3026 = 3.5373$$

$$C = 1.0463$$

$$\chi^2_{\text{corr.}} = \frac{3.5373}{1.0463} = 3.38 \text{ with 3 df}$$

$$\chi^2_{.05, 3 \text{ df}} = 7.81$$

Since $3.38 < 7.81$ accept the hypothesis that the variance are equal at 0.05 α -level.

Calculations for Sum of Squares
(Bulk Specific Gravity for C 128-57)

axb Table

Operator	Material		Totals
	FM 2.40	FM 3.00	
A	26.15	25.98	52.13
B	26.03	25.88	51.91
Totals	52.18	51.86	104.04

$$\frac{T^2}{rab} = \frac{(104.04)^2}{40} = \frac{10824.3216}{40} = 270.60804$$

$$\sum \sum Y_{ijk}^2 = 270.6158$$

$$\frac{\sum R_j^2}{ab} = \frac{1082.4362}{4} = 270.60905$$

$$\frac{\sum T_{jk}^2}{r} = 270.6118$$

$$\frac{\sum A_j^2}{rb} = \frac{5412.2120}{20} = 270.6106$$

$$\frac{\sum B_k^2}{ra} = \frac{5412.1850}{20} = 270.60925$$

$$\Sigma_y^2 = 270.6158 - 270.6080 = 0.0078$$

$$R_{yy} = 270.6090 - 270.6080 = 0.0010$$

$$S_{ab} = 270.6118 - 270.6080 = 0.0038$$

$$A_{yy} = 270.6106 - 270.6080 = 0.0026$$

$$B_{yy} = 270.6092 - 270.6080 = 0.0012$$

$$(AB)_{yy} = S_{ab} - A_{yy} - B_{yy} = 0$$

$$E_{yy} = \Sigma_y^2 - R_{yy} - S_{ab} = 0.0030$$



Tests of Hypotheses
(Bulk Specific Gravity for C 128-57)

$H_1: \sigma_r^2 = 0$ (Replicate effect)

$$\frac{MS_r}{MS_E} = \frac{0.0001}{0.0001} = 1 < 2.25 = F_{.05} \quad 9, 27df$$

Therefore accept H_1 .

$H_2: \sigma_{\alpha\beta}^2 = 0$ (Operator-Material Interaction effect)

$$\frac{MS_{\alpha\beta}}{MS_E} = \frac{0}{0.0001} = 0 < 4.21 = F_{.05} \quad 1, 27df$$

Therefore accept H_2 .

Pooling: $SS_E + SS_r + SS_{\alpha\beta} = 0.0040$

$$df_E + df_r + df_{\alpha\beta} = 37$$

$$S^2_{\text{pooled}} = \frac{0.0040}{37} = 0.0001 = MS_{\text{pooled}}$$

$H_3: \sigma_\alpha^2 = 0$ (Material effect)

$$\frac{MS_\alpha}{MS_{\text{pooled}}} = \frac{0.0026}{0.0001} = 26 > 4.10 = F_{.05}, \quad 1, 37df$$

Therefore we are unable to accept H_3 .

$H_4: \sigma_\beta^2 = 0$ (Operator effect)

$$\frac{MS_\beta}{MS_{\text{pooled}}} = \frac{0.0012}{0.0001} = 12 > 4.10 = F_{.05}, \quad 1, 37df$$

Therefore we are unable to accept H_4 .

Conclusions: Operator and materials are significant. Their interactions and replicates are not significant.

The same conclusions may be made for Bulk (SSD), however, there are no significant variance components in the analysis of apparent sp. gr.

Bartlett Test for Homogeneity of Variance
(Bulk SSD Specific Gravity for C 128-57)

	FM = 2.40		FM = 3.00	
	Operator A	Operator B	Operator A	Operator B
X.	26.40	26.31	26.33	26.23
\bar{X}	2.640	2.631	2.633	2.623
$\sum X_i^2$	69.6966	69.2219	69.3277	68.8019
$\frac{\sum X_i^2}{r}$	69.6960	69.2216	69.3269	68.8013
SS	0.0006	0.0003	0.0008	0.0006
df	9	9	9	9
S_j^2	0.00007	0.00003	0.00009	0.00007
$\log S_j^2$	- 4.15490	- 4.52288	- 4.04576	- 4.15490
$(r-1) \log S_j^2$	-37.39410	-40.70592	-36.41184	-37.39410

$$\sum (r-1) \log S_j^2 = - 151.90596$$

$$\sum SS = 0.00026 \quad S_j^2 = \frac{0.00026}{36} = 0.00007$$

$$\log S_j^2 = - 4.15490 \times 36 = -149.57640$$

$$2.32956 \times 2.3026 = 5.3640$$

$$C = 1.0463$$

$$\chi^2_{\text{corr.}} = \frac{5.3640}{1.0463} = 5.13 \text{ with 3df}$$

$$\chi^2_{.05, 3df} = 7.81$$

Since 5.13 < 7.81, accept the hypothesis that the variances are equal at 0.05 α -level.

Calculations for Sum of Squares
(Bulk SSD Specific Gravity for C 128-57)

axb Table

Operator	Material		Totals
	FM 2.40	FM 3.00	
A	26.40	26.33	52.73
B	26.31	26.23	52.54
Totals	52.71	52.56	105.27

$$\frac{T^2}{rab} = \frac{(105.27)^2}{40} = \frac{11081.7729}{40} = 277.0443$$

$$\Sigma \Sigma \Sigma Y_{ijk}^2 = 277.0481$$

$$\frac{\Sigma R_i^2}{ab} = \frac{1108.1803}{4} = 277.0451$$

$$\frac{\Sigma T_{jk}^2}{r} = 27.0458$$

$$\frac{\Sigma A_j^2}{rb} = \frac{5540.8977}{20} = 277.0449$$

$$\frac{\Sigma B_k^2}{ra} = \frac{5540.9045}{20} = 277.9452$$

$$E_y^2 = 277.0481 - 277.0443 = 0.0038$$

$$R_{yy} = 277.0451 - 277.0443 = 0.0008$$

$$S_{ab} = 277.0458 - 277.0443 = 0.0015$$

$$A_{yy} = 277.0449 - 277.0443 = 0.0006$$

$$B_{yy} = 277.0452 - 277.0443 = 0.0009$$

$$(AB)_{yy} = S_{ab} - A_{yy} - B_{yy} = 0$$

$$E_{yy} = \Sigma_y^2 - R_{yy} - S_{ab} = 0.0015$$

Bartlett Test for Homogeneity of Variance
(Apparent Specific Gravity for C 128-57)

	FM = 2.40		FM = 3.00	
	Operator A	Operator B	Operator A	Operator B
X	26.81	26.79	26.81	26.83
\bar{X}	2.681	2.679	2.681	2.683
$\sum X_i^2$	71.8787	71.7711	71.8783	71.9853
$\frac{\sum x_i^2}{r}$	71.8776	71.7704	71.8776	71.9849
SS	0.0011	0.0007	0.0007	0.0004
df	9	9	9	9
S_j^2	0.00012	0.00008	0.00008	0.00004
$\log S_j^2$	- 3.92082	- 4.09691	- 4.09691	- 4.39794
$(r-1) \log S_j^2$	-35.28728	-36.87219	-36.87219	-39.58146

$$\sum (r-1) \log S_j^2 = - 148.61322$$

$$\sum SS = 0.00032 \quad \frac{0.00032}{30} = 0.00009$$

$$\log S_j^2 = - 4.04576 \times 36 = - 145.64736$$

$$2.96586 \times 2.3026 = 6.8292$$

$$C = 1.0463$$

$$\chi^2_{\text{corr.}} = \frac{6.8292}{1.0463} = 5.53 \text{ with 3df}$$

$$\chi^2_{.05, 3df} = 7.81$$

Since $5.53 < 7.81$ accept the hypothesis that the variances are equal at 0.05 α -level.

Calculations for Sum of Squares
(Apparent Specific Gravity for C 128-57)

Operators	Material		Totals
	FM = 2.40	FM = 3.00	
A	26.81	26.81	53.62
B	26.79	26.83	53.62
Totals	53.60	53.64	107.24

$$\frac{T^2}{rab} = \frac{(107.24)^2}{40} = \frac{11500.4176}{40} = 287.5104$$

$$\Sigma \Sigma Y^2_{ijk} = 287.5134$$

$$\frac{\Sigma R_i^2}{ab} = \frac{1150.0454}{4} = 287.5114$$

$$\frac{\Sigma T_{jk}}{r} = 287.5105$$

$$\frac{\Sigma A_j^2}{rb} = \frac{5750.2096}{20} = 287.5105$$

$$\frac{\Sigma B_k^2}{ra} = \frac{5750.2088}{20} = 287.5104$$

$$\Sigma_y^2 = 287.5134 - 287.5104 = 0.0030$$

$$R_{yy} = 287.5114 - 287.5104 = 0.0010$$

$$S_{ab} = 287.5105 - 287.5104 = 0.0001$$

$$A_{yy} = 287.5105 - 287.5104 = 0.0001$$

$$B_{yy} = 287.5104 - 287.5104 = 0$$

$$(AB)_{yy} = S_{ab} - A_{yy} - B_{yy} = 0$$

$$E_{yy} = \Sigma_y^2 - R_{yy} - S_{ab} = 0.0019$$

